Synthesis Lectures on
Mathematics & Statistics

SYNTHESIS
COLLECTION OF TECHNOLOGY

William C. Bauldry

Computational Calculus

A Numerical Companion to Elementary Calculus

Springer

Synthesis Lectures on Mathematics & Statistics

Series Editor

Steven G. Krantz, Department of Mathematics, Washington University, Saint Louis, MO, USA

This series includes titles in applied mathematics and statistics for cross-disciplinary STEM professionals, educators, researchers, and students. The series focuses on new and traditional techniques to develop mathematical knowledge and skills, an understanding of core mathematical reasoning, and the ability to utilize data in specific applications.

William C. Bauldry

Computational Calculus

A Numerical Companion to Elementary Calculus

Springer

William C. Bauldry
Appalachian State University
Boone, NC, USA

ISSN 1938-1743 ISSN 1938-1751 (electronic)
Synthesis Lectures on Mathematics & Statistics
ISBN 978-3-031-29657-4 ISBN 978-3-031-29658-1 (eBook)
https://doi.org/10.1007/978-3-031-29658-1

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Dedicated to the memory of my wife,
Sue Ann Bauldry, Ph.D.
Fifty-one years together were too few.

Preface

Dividing one number by another is mere computation; knowing what to divide by what is mathematics.

Jordan Ellenberg, How Not to Be Wrong: The Power of Mathematical Thinking.

This monograph is intended to supplement an introductory calculus course. The text is self-contained as much as possible so as to be useable with any standard calculus textbook. Most popular calculus textbooks have very little material covering numerical integration and even less on numerical differentiation. A large percentage of the students in our calculus courses are not mathematics majors. These students will rarely see a function specified by a formula after they leave the course and won't have the tools to be able to apply the incredibly powerful calculus methods they've learned in their subsequent work. This text hopes to help ameliorate that situation by introducing computational methods for elementary calculus supported by using a computer algebra system (CAS). The CAS we'll use is Maple™; however, any CAS can be used with minor, mostly syntactic, modifications.

The text begins with the Taylor polynomials and developing difference formulas for derivatives. Error analysis is introduced early and is part of each formula's development. Chapter 2 investigates numerical integration from the traditional calculus course viewpoint. Again error analysis is introduced early and stays a part of each method's development. The Gaussian and Gauss-Kronrod quadratures are studied with the tie that these methods are the standard for Texas Instruments' calculators. As the error bounds involve deeper mathematics for their development, the bounds are just stated for these quadratures. The last chapter is a collection of 15 student projects that can be used as group work.

Students have enjoyed the study and have found that it enhances their understanding of calculus. And that, along with being useful, can be great fun.

To the Student

We spend most of our effort when studying elementary calculus in learning to find derivatives and integrals algebraically. We quickly see that differentiation is much simpler than integration. For derivatives, we learn a small set of rules like the "Power Rule" and the "Product Rule" that, together with examples, can fit on a ASU.[1] Nevertheless, functions given as data or by procedures can be quite difficult to differentiate—they don't fit the algebraic rules. Then we discover that, in general, integrals are much more difficult to find than derivatives. Most calculus texts contain tables of standard forms of integrals for using pattern matching with substitution as the main technique. We also learn a small collection of heuristics like "LIATE" (*Logarithmic*, *Inverse Trigonometric*, *Algebraic*, *Trigonometric*, *Exponential*) to help with integration by parts. Unfortunately, in contrast to a derivative, a typical function's integral cannot be expressed in a simple form. The first such *special function* we encounter in calculus is the integral of $1/t$ from 1 to x, defined as the natural logarithm $\ln(x)$. In the past, students used a table or a slide rule to compute a logarithm; now we use a scientific calculator or a computer. Even though most functions do not have an elementary antiderivative, we can approximate nearly all definite integrals using numerical techniques.

But howcan the values of derivatives and definite integrals actually be calculated? That is the subject of this monograph.

Boone, NC, USA
February 2023

William C. Bauldry

Acknowledgements Thanks go to Susanne Filler, an excellent editor, and her team, especially to Boopalan Renu and Vidya Velmurugan who made production go so easily.

A Note on Computer Algebra Systems

We will use the computer algebra system[2] Maple™ from Maplesoft™ throughout this monograph. Maple can be found at www.maplesoft.com/products/maple/. A Maple session can be in either of two modes: Document and Worksheet. While Document mode is more flexible incorporating more text editing capability, we will use Worksheet mode as input follows a prompt symbol, >, and the mode displays cell markers by default making the code easier for students to parse. The version used in the text is Maple 2023; however, versions back to 2016 will have no issues.

[1] Dark blue text indicates a web-link. Cyan text indicates a link within this text.

[2] All product and/or company names shown which are trademarks or registered trademarks of their respective holders do not imply any affiliation with or endorsement by them.

The Maple Calculator™ (free for iOS and Android; information is available at www.maplesoft.com/products/Maplecalculator/) can take a photo of mathematics in a text or that is handwritten, operate on the captured expressions, or transfer them to desktop Maple through the Maple Cloud™ (maple.cloud). Maple Cloud is a convenient repository of Maple documents that is publicly accessible and easy to use. An instructor can distribute documents to students quickly through the Maple Cloud.

Along with Maple and Maple Calculator, Maplesoft also offers Maple Player™, "a free application that lets you view and even interact with Maple documents to solve problems, visualize solutions, and explore concepts when you don't have Maple." With careful design, an instructor can make interactive worksheets to provide students with a microworld for investigating concepts with Maple Player. For more information, see www.maplesoft.com/products/maple/Mapleplayer/.

Current Computer Algebra Systems

The Maple examples, problems, etc., can easily be implemented in any CAS.[3]

- Axiom
  sourceforge.net/projects/axiom/
- FriCAS
  fricas.github.io
- Mathematica
  www.wolfram.com/mathematica/
- Mathics
  mathics.org
- Maxima
  maxima.sourceforge.io
- PocketCAS (macOS™ and iOS™)
  pocketcas.com
- Reduce
  reduce-algebra.sourceforge.io
- SageMath
  www.sagemath.org
  CoCalc Portal for web-based SageMath
  cocalc.com/?utm source=sagemath.org
- Symbolic Math Toolbox (formerly MuPAD) for MATLAB
  www.mathworks.com/discovery/mupad.html

[3] Axiom, FriCAS, Maxima, Mathics, Reduce, SageMath, Xcas, and Yacas are free.Wolfram Alpha has free and paid portals. Costs of the other CASs can be found on their respective websites. Most colleges and universities hold site licenses for one or more of the commercial systems.

- Wolfram Alpha, a web-based (resource limited) access to Mathematica
  www.wolframalpha.com
- Xcas
  https://xcas.univ-grenoble-alpes.fr/en.html
  web-based Xcas:
  https://xcas.univ-grenoble-alpes.fr/xcasjs/
- Yacas
  www.yacas.org

Symbolic computations can be programmed with SymPy, the symbolic mathematics library for Python; see www.sympy.org/en/index.html.

Calculators with Computer Algebra Capabilities

The Maple examples, problems, etc., can also be done with a CAS calculator.

- Casio CAS Graphing Scientific Calculator
  edu.casio.com/products/cg/cp2/
- Hewlett-Packard HP Prime Graphic Calculator
  hpofficesupply.com/product/hp-prime-graphing-calculator/
- TI-Nspire CAS calculators
  education.ti.com/en/products/calculators/.

Contents

1 Numerical Differentiation 1
- 1.1 Taylor's Theorem 2
  - 1.1.1 Forms of the Remainder 5
  - 1.1.2 "Big $\mathcal{O}$" Notation 6
- 1.2 Proving Taylor's Theorem 7
- 1.3 Maple and Taylor Expansions 8
- 1.4 Tailored Expressions for the First Derivative 10
  - 1.4.1 Two Useful Taylor Expansions 11
  - 1.4.2 Forward Difference Approximation 11
  - 1.4.3 Backward Difference Approximation 14
  - 1.4.4 Centered Difference Approximation 16
  - 1.4.5 Comparing the Error Terms 21
- 1.5 Higher Derivatives 22
- 1.6 Next Steps 24
- 1.7 Taylor's Statement of the Theorem 27
- 1.8 Centered Difference Coefficients Chart 28
- 1.9 Links and Others 29

References 30

2 Numerical Integration 31
- 2.1 Methods of Elementary Calculus 31
  - 2.1.1 Rectangle Methods 32
  - 2.1.2 Trapezoid Sums 43
  - 2.1.3 Simpson's Rule 47
  - 2.1.4 A Maple Comparison 50
  - 2.1.5 A Chart of the Methods 51
- 2.2 The Basic Integration Methods and Maple 52
- 2.3 Gaussian Quadrature 54
  - 2.3.1 In Search of Improvements 55
  - 2.3.2 Patterns 55

2.3.3 Testing Gauss 57
2.3.4 An Error Bound for Gaussian Quadrature 58
2.3.5 Values of Gaussian Weights and Nodes 59
2.4 Gauss-Kronrod Quadrature 60
2.4.1 Aleksandr Kronrod's Idea 60
2.4.2 Gauss-Kronrod Quadrature in Practice 61
2.5 Transforming the Integral Over $[a, b]$ to $[-1, 1]$ 63
2.6 Easy, but Hard: A Class Quadrature Project 64
2.7 Next Steps 65
2.8 A Menagerie of Test Integrals 69
2.9 Legendre and Stieltjes Polynomials for $GK_{7,15}$ 70
2.10 Links and Others 71
References 72

3 Projects 73
3.1 Polynomials and Roots from Data Using Numerical Derivatives 73
3.1.1 Monic Polynomials and Their Derivatives 73
3.1.2 The Roots 74
3.1.3 The Project 74
3.2 Space Shuttle Acceleration 75
3.2.1 Space Shuttle Acceleration: The Situation 75
3.2.2 The Project 76
3.3 Measuring Capacitance and Inductance with Derivatives 76
3.3.1 Capacitance and Inductance Equations 76
3.3.2 The Project: Measuring Component Values 77
3.4 Marginal Revenue From Data, Cost, and Profit 77
3.4.1 Marginal Revenue and Marginal Cost 77
3.4.2 The Project 78
3.5 Estimating Speed and Acceleration from GPS Data 80
3.5.1 The Global Positioning System and Distance 80
3.5.2 The Project 80
3.6 Fornberg's Algorithm 82
3.6.1 The Algorithm 82
3.6.2 The Project 83
3.7 Euler's Method of Solving Differential Equations 83
3.7.1 Euler's Method Graphically 83
3.7.2 Euler's Method Algebraically 84
3.7.3 The Project 84
3.8 A Dog's Life 86
3.8.1 Rate of Aging Data 86
3.8.2 The Project 87

3.9 Computing the Cost at a Gas Pump 87
3.9.1 Calculating the Cost of the Gas Pumped 87
3.9.2 The Project 88
3.10 The Fourier Power Spectrum 89
3.10.1 Calculating the Power Spectrum of a Signal 90
3.10.2 The Project: Produce a Spectrum Plot of the Square Wave 90
3.10.3 Extra for Experts 90
3.11 One Function for All, The Normal Distribution 91
3.11.1 The Project 92
3.12 Monte Carlo Integration 92
3.12.1 Algebraic: The Definite Integral From an Average 93
3.12.2 Geometric: The Definite Integral From Areas 94
3.12.3 The Project 96
3.13 Computing a Transition Curve 97
3.13.1 Transition Curves 97
3.13.2 The Project: Computing a Transition Curve for a Roadbed 97
3.14 Finding the Factorial of $-1/2$ 98
3.14.1 The Gamma Function 98
3.14.2 The Project 99
3.15 Theory: Plugging the Hole in the “Power Rule” 100
3.15.1 The Integral “Power Rule” 100
3.15.2 The Project 100
References 101

Further Readings 103

Index 105

Numerical Differentiation 1

What is *Numerical Differentiation?* Numerical Differentiation is the approximation of the derivative of a function at a point using numerical formulas, not the algebraic rules or the limit definition for differentiation. The basic form uses the slope of a short chord rather than the tangent line. Since we may be subtracting numbers that are close together to compute the chord's slope, loss of computational precision can be a serious problem. More advanced forms of estimating a derivative use weighted combinations of the function's values at nearby points (Fig. 1.1).

There are many reasons for finding a derivative numerically. We often may have data rather than a formula for the function; the function may be specified by a set of steps or an algorithm not giving a simple formula; or the function's formula may be very complex and 'expensive' to compute.

Taylor series expansions will be our basic tool to develop formulas and error bounds for numerical derivatives. The errors will have two main components: *truncation errors* from using Taylor polynomials rather than the function's full infinite series, and *round-off errors* that are inherent from using finite-precision floating-point arithmetic.

A simple example of truncation error is using $y = x$ in place of $\sin(x)$ for small values of x. At $x = 0$, the two functions are equal, but as x moves away from zero, the truncation error gets larger and larger. *Graph* $x - \sin(x)$ over $[-\pi, \pi]$. This example occurs in analyzing the model of an oscillating simple pendulum: $\theta'' + \alpha \sin(\theta) = 0$.

A simple example of round-off error is shown on a TI-84 series calculator. A calculator uses finite precision arithmetic so numbers like $\frac{1}{3}$ cannot be expressed exactly. Enter: `1+10^-13` → `X`. Now compare the answers to `X-1` and `X-0.9-0.1`. They're different! The problem is due to rounding error. Think about how this problem could affect a bank.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

W. C. Bauldry, *Computational Calculus*, Synthesis Lectures on Mathematics & Statistics,
https://doi.org/10.1007/978-3-031-29658-1_1

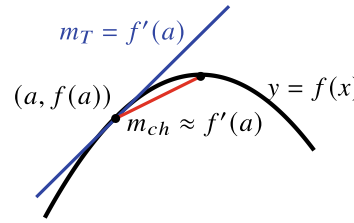


Fig. 1.1 Approximating a derivative with a chord

1.1 Taylor's Theorem

Brook Taylor developed the polynomials that are named for him, writing[1] about them in his 1715 text *Methodus Incrementorum Directa et Inversa*.[2] Taylor started with Newton's polynomial interpolation and recognized that divided differences approached a function's derivatives as the step-size went to zero. Mathematicians thought, for a number of years, that this calculation worked for every function.

Recall that we compute a function's Taylor polynomial by matching its derivatives up to the order equalling the degree of the polynomial. Davis, Porta, and Uhl called this "matching order of contact."[3] Building a polynomial using order of contact matches the function and its behavior near a point a; it's a local phenomenon just like derivative. The Taylor polynomial matches the function's value (*order* 0), matches the function's slope (*order* 1), matches the function's concavity (*order* 2), matches the function's change in concavity (*order* 3), etc. When the point of contact is $a = 0$, the polynomial is called a Maclaurin polynomial.

Definition 1.1 *(Taylor Polynomials (1712))* If f has sufficiently many derivatives at $x = a$, the *Taylor polynomial of degree n* (or order n) is

$$p_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} (x - a)^k$$

where we choose $f^{(0)}(a) = f(a)$.

Let's look at examples of Taylor polynomials for a set of important functions. *Work through the details of each example!*

[1] James Gregory had discovered several "Taylor polynomials" in 1671, 14 *years before Taylor was born!* In 1664, before Newton investigated calculus, Gregory studied "the twin keys to the calculus, the method of tangents (differentiation) and of quadratures (integration)" under Angeli in Italy.

[2] See *Methodus Incrementorum Directa et Inversa*, Londini, 1717, or Ian Bruce's annotated translation.

[3] See Davis, Porta, and Uhl, *Calculus & Mathematica*, [1].

Selected Taylor Polynomials.

In each of the example Taylor polynomials below, identify the pattern in the new terms and predict what the next higher order Taylor polynomial would be.

1. Let $f(x) = \cos(x)$ and $a = 0$. For $n = 4$ to 10,

$$p_4(x) = p_5(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4,$$

$$p_6(x) = p_7(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6,$$

$$p_8(x) = p_9(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \frac{1}{8!}x^8,$$

$$p_{10}(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \frac{1}{8!}x^8 - \frac{1}{10!}x^{10}.$$

2. Let $g(x) = e^x$ and $a = 0$. For $n = 4$, 5, and 6,

$$p_4(x) = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4,$$

$$p_5(x) = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \frac{1}{5!}x^5,$$

$$p_6(x) = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \frac{1}{5!}x^5 + \frac{1}{6!}x^6.$$

3. Let $h(x) = \ln(1 + x)$ and $a = 0$. For $n = 3$, 4, 5, and 6,

$$p_3(x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3.$$

$$p_4(x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4,$$

$$p_5(x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \frac{1}{5}x^5,$$

$$p_6(x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \frac{1}{5}x^5 - \frac{1}{6}x^6.$$

4. Let $k(x) = \arctan(x)$ and $a = 0$. For $n = 5$ to 10,

$$p_5(x) = p_6(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5,$$

$$p_7(x) = p_8(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7,$$

$$p_9(x) = p_{10}(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \frac{1}{9}x^9.$$

5. Let $j(x) = 1/(1-x)$ and $n = 4$, 5, and 6,

   a. For $a = 0$,

$$p_4(x) = 1 + x + x^2 + x^3 + x^4,$$

$$p_5(x) = 1 + x + x^2 + x^3 + x^4 + x^5,$$

$$p_6(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6.$$

   b. For $a = 1/2$,

$$p_4(x) = 2 + 2^2\left(x - \tfrac{1}{2}\right) + 2^3\left(x - \tfrac{1}{2}\right)^2 + 2^4\left(x - \tfrac{1}{2}\right)^3 + 2^5\left(x - \tfrac{1}{2}\right)^4,$$

$$p_5(x) = 2 + 2^2\left(x - \tfrac{1}{2}\right) + 2^3\left(x - \tfrac{1}{2}\right)^2 + 2^4\left(x - \tfrac{1}{2}\right)^3 + 2^5\left(x - \tfrac{1}{2}\right)^4 + 2^6\left(x - \tfrac{1}{2}\right)^5,$$

$$p_6(x) = 2 + 2^2\left(x - \tfrac{1}{2}\right) + 2^3\left(x - \tfrac{1}{2}\right)^2 + 2^4\left(x - \tfrac{1}{2}\right)^3 + 2^5\left(x - \tfrac{1}{2}\right)^4 + 2^6\left(x - \tfrac{1}{2}\right)^5 + 2^7\left(x - \tfrac{1}{2}\right)^6.$$

   c. For $a = 1$. There are no Taylor polynomials for $j(x)$ at $a = 1$. *Explain why!*

To compute a function's Taylor polynomial, we must calculate its derivatives at a point. Often this calculation is easy like those for $f(x) = \cos(x)$ and $g(x) = e^x$ in the examples above. Sometimes, however, it's quite difficult to see a pattern. Look at the first three derivatives of $f(x) = \tan(\sin(x))$. *Verify these with Maple!*.

$$\begin{aligned}
f(x) &= \tan(\sin(x)) \\
f'(x) &= \cos(x) \cdot \sec^2(\sin(x)) \\
f''(x) &= \sec^2(\sin(x)) \cdot \left(2\cos^2(x)\tan(\sin(x)) - \sin(x)\right) \\
f'''(x) &= \cos(x)\sec^2(\sin(x))\left(6\sec^2(\sin(x)) - 4\cos^2(x) - 6\sin(x)\tan(\sin(x)) - 1\right)
\end{aligned}$$

Does there appear to be a formula or a pattern?

How well does a Taylor polynomial fit a function? Taylor's theorem tells us how good an approximation the polynomial is by explicitly giving the remainder R_n; that is, by explicitly giving the error.

Theorem 1.1 (Taylor's Theorem) *Suppose f has $n+1$ derivatives on an open interval containing a. Then $f(x) = p_n(x) + R_n(x)$ where*

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

for some $c = c(a,x)$ between x and a. We call R_n the remainder.

The notation $c = c(a,x)$ indicates that the value of c depends on both x and a.

1.1.1 Forms of the Remainder

If we are able to analyze the error term R_n, then we can know how well the Taylor polynomial represents a function near the point $x = a$. Table 1.1 lists four forms of R_n that have been developed; each has advantages in different situations. The uniform form will be[4] our tool of choice since it will be the simplest to work with as we won't need to find an intermediate point c or evaluate an integral. Since the uniform form uses a maximum value, $M_{n+1} \geq \max\left|f^{(n+1)}(x)\right|$, it tends to provide more conservative estimates than necessary, but the ease of use makes it worthwhile.

Lagrange appears to have been the first to study what conditions are needed for a function to have a useful Taylor series expansion at a point a; i.e., that the remainder must go to zero as n increases for all x values in $|x-a| < r$ for some $r > 0$. A function that has a series expansion that converges on an interval is called *analytic*.

Table 1.1 Forms of the Remainder R_n

Lagrange (1797)	$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$
Cauchy (1821)	$R_n(x) = \frac{f^{(n+1)}(c)}{n!}(x-c)^n(x-a)$
Integral form	$R_n(x) = \frac{1}{n!}\int_a^x f^{(n+1)}(t)\,(x-t)^n\,dt$
Uniform form	$\lvert R_n(x)\rvert \leq \frac{\lvert x-a\rvert^{n+1}}{(n+1)!}\cdot M_{n+1}$ where $M_{n+1} \geq \max\left\lvert f^{(n+1)}(x)\right\rvert$

[4] Cauchy also discovered the remainder's Integral form in 1821. See [2], pg. 116.

1.1.2 "Big $\mathcal{O}$" Notation

Paul Bachmann and Edmund Landau's "Big $\mathcal{O}$" notation let's us easily express the size of errors. The notation is sometimes "asymptotic notation." The formal definition is

Definition 1.2 *(Landau Notation)* Let f and g be two functions. Then we say $f(x) = \mathcal{O}(g(x))$ as $x \to \infty$ if and only if there is a positive constant c and a value x_0 such that

$$|f(x)| \leq c \cdot g(x) \text{ for all } x \geq x_0.$$

At a point $x = a$, we have $f(x) = \mathcal{O}(g(x))$ as $x \to a$ if and only if there are positive constants c and δ such that

$$|f(x)| \leq c \cdot g(x) \text{ when } 0 < |x - a| < \delta.$$

When the context is clear, we can write $f = \mathcal{O}(g)$.

In words, Big $\mathcal{O}$ notation says

- For x going to infinity, the magnitude of $f(x)$ is *about* (less than a constant times) the magnitude of $g(x)$ when x is *large enough*.
- For x near a, the magnitude of $f(x)$ is *about* (less than a constant times) the magnitude of $g(x)$ when x is *close enough* to a.

Big $\mathcal{O}$ Notation Examples

1. Let y be any parabola. Then $y = \mathcal{O}(x^2)$ since $|ax^2 + bx + c| \leq Ax^2$ for $|x| > 1$ for some A, for example, $A = |a| + |b| + |c|$ will work.
2. Let $z = \sin(x)$ Then
   a. $z = \mathcal{O}(1)$ as $x \to \infty$, and
   b. $z = \mathcal{O}(x)$ as $x \to 0$.
3. The function $f(x) = e^x$ is *not* $\mathcal{O}(x^n)$ for any integer n since $\lim_{x\to\infty} x^n/e^x = 0$ for any $n \in \mathbb{N}$. On the other hand, $x^n = \mathcal{O}(e^x)$ for every $n \in \mathbb{N}$.
4. The natural logarithm $\ln(x) = \mathcal{O}(x^{1/n})$ for any $n \in \mathbb{N}$. L'Hôpital's rule easily shows this result.

Using Big $\mathcal{O}$ will let us write the uniform form of the remainder very simply and concisely as

$$|R_n(x)| = O\left(|x-a|^{n+1}\right),$$

or, with $h = |x-a|$,

$$|R_n(x)| = O\left(h^{n+1}\right).$$

There is an arithmetic of Big $\mathcal{O}$ terms. Adding two functions: $\mathcal{O}(f_1) + O(f_2) = O(\max\{f_1, f_2\})$. For instance, $\mathcal{O}(x^2) + O(x^3) = O(x^3)$ for x getting large, while $\mathcal{O}(h^2) + O(h^3) = O(h^2)$ for h going to 0. Multiplying by a big O term by constant is quite easy: $c \cdot O(g) = O(g)$. For example, $100 \cdot O(h^2) = O(h^2)$. Lastly, divide a big O term by a power: $\mathcal{O}(h^n)/h^p = O(h^{n-p})$; e.g., $\mathcal{O}(h^5)/h^2 = O(h^3)$. *Explain why these are all true!*

For more information about Landau notation, start with "Asymptotic Notation."

1.2 Proving Taylor's Theorem

For a brief side trip, we'll sketch an outline of a proof of Taylor's Theorem—it's a very nice application of integration by parts.

Proof (Taylor's Theorem — *Outline*.)

1. The Fundamental Theorem of Calculus gives $f(x) = f(a) + \int_0^{x-a} f'(x-t)\,dt$.
2. Integrate by parts with $u = f'(x-t)$ and $dv = dt$ to obtain

$$f(x) = f(a) + f'(a)(x-a) + \int_0^{x-a} f''(x-t)\cdot t\,dt$$

3. Repeat the process: choosing $u = f^{(k)}(x-t)$ and $dv = t^{k-1}/(k-1)!$ successively to arrive at

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(a,x)$$

where

$$R_n(a,x) = \frac{1}{n!}\int_0^{x-a} f^{(n+1)}(x-t)\cdot t^n\,dt$$

Recall that the notation $R_n(a,x)$ indicates that both a and x determine the value of the remainder.

A simple substitution in the remainder above gives the Integral Form from Table 1.1. To obtain the form shown in Taylor's theorem, Theorem 1.1, apply the *Mean Value Theorem for Definite Integrals* to the remainder above.

The theory of Taylor polynomials and series is both important and fascinating. To fully understand Taylor series requires complex variables, but we can use them quite profitably staying in the real domain. See [3] to begin further study. To see the role of Taylor's work in the development of calculus, start with [4].

1.3 Maple and Taylor Expansions

Computer algebra systems can compute Taylor polynomials symbolically for a vast number of functions very quickly. The two Maple functions that we will use the most are *taylor* and *mtaylor*.[5] The syntax of *taylor* is

```
> taylor(expression, variable = base, order);
```

To use the default base of $a = 0$ and the default order 6, just enter:

```
> taylor(expression, variable);
```

We can evaluate Maple's Taylor expansion at $x = a$ by using *eval* as in

```
> T := taylor(expression, variable);
  eval(T, x = a);
```

For example, let's use Maple to compute the Taylor expansion for $f(x) = \ln(1 - x)$ at $x = 0$ to order $\mathcal{O}(x^6)$, and evaluate the result at $x = 1$.

A function can fail to have a Taylor polynomial at a point for a variety of reasons. For example, $A(x) = |x|$ has no Taylor polynomials based at $x = 0$. *Explain why!* Maple will just give an error for *taylor*$(|x|, x)$ that says "no series at 0."

Computing Taylor Approximations with Maple

Remember the syntax for *taylor:* the expression, then the variable.

```
> T := taylor(ln(1 + x), x)
```

$$T := x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \frac{1}{5}x^5 + \mathrm{O}\left(x^6\right)$$

We can evaluate expressions containing the result with *eval.*

```
> eval(ln(1 + x), x = 1), eval(ln(1 + x), x = 1.0);
  eval(T, x = 1), eval(T, x = 1.0);
```

$$ln(2),\ 0.6931471806$$

$$\frac{47}{60},\quad 0.7833333333$$

Note the two different forms above: 1 and 1.0; the decimal forces "floating point" arithmetic; i.e., decimal approximation.

[5] Maple has a more general command *series* that will compute Taylor expansions when possible.

Often functional notation is a better way to proceed. To be able to use *unapply* to define a function from the expression, we must first *convert* the series to a polynomial removing the Big $\mathcal{O}$ term.

```
> p := convert(T, polynom);
  P := unapply(p, x);
```

$$x - \frac{1}{2}\cdot x^2 + \frac{1}{3}\cdot x^3 - \frac{1}{4}\cdot x^4 + \frac{1}{5}\cdot x^5$$

$$P := x \mapsto x - \frac{1}{2}\cdot x^2 + \frac{1}{3}\cdot x^3 - \frac{1}{4}\cdot x^4 + \frac{1}{5}\cdot x^5$$

```
> P(1), P(1.0);
```

$$\frac{47}{60},\ 0.7833333333$$

The *mtaylor* function is defined to compute multivariable Taylor polynomials. We'll use it to produce a polynomial, avoiding a series with a Big $\mathcal{O}$ term. For our use, the syntax is the same as *taylor.*

```
> mtaylor(ln(1 + x), x);
```

$$x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \frac{1}{5}x^5$$

There are two very useful routines for understanding Taylor polynomials that are found in the *Calculus1* subpackage of the *Student* package: *TaylorApproximationTutor* and *TaylorApproximation.* To make these programs available, use *with* to load the package; i.e., execute

```
> with(Student[Calculus1]) :
```

Figure 1.2 shows the window generated by

```
> TaylorApproximationTutor(ln(1 + x));
```

Maple also includes a "Math App" for investigating Taylor expansions. Select *Math Apps* from the *Tools* menu, then double-click the *Taylor's Theorem* app in the *Calculus* section.

We close this section with a class activity using the *TaylorApproximationTutor*.

"Think and Share" Activity

Small Groups:

Load the *TaylorApproximationTutor* program by executing

```
> with(Student[Calculus1]) :
```

Now define the function $f(x) = \tan(\sin(x))$ and start the Taylor Approximation Tutor with

```
> f := x → tan(sin(x)) :
  TaylorApproximationTutor(f(x));
```

Change the plot window to $[-2\pi, 2\pi] \times [-3, 3]$, and then create an animation of the Taylor

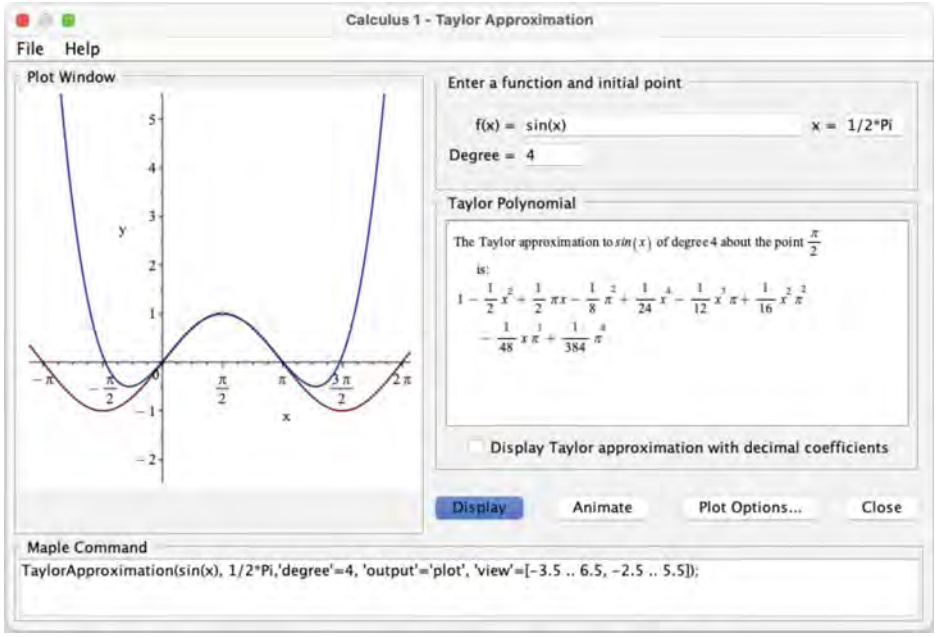


Fig. 1.2 *TaylorApproximationTutor*'s Window

polynomials of the function. Observe how well and where the polynomials match the function as the degree increases.

Class: Compare the groups' observations and discuss potential reasons for the behavior observed.

1.4 Tailored Expressions for the First Derivative

We have all the machinery we need to develop good formulas with error estimates for approximating a function's derivative at a point. We'll use generic Taylor polynomials to build our approximation formulas, not the specific forms that would require us to compute a number of the function's derivatives—the very values we want to approximate! As we develop the formulas, watch to see why a simple two-point slope formula would prove to be inadequate.

1.4.1 Two Useful Taylor Expansions

Two special Taylor polynomials will lead to our formulas for approximating the derivative of a function at a point. Start by "looking forward" from $x = a$ to $x = a + h$; then by "looking backward" from $x = a$ to $x = a - h$ The Big-$\mathcal{O}$ terms will make it easy for us to keep track of error estimates.

First, the forward direction. Set $x = a + h$ in the general Taylor polynomial. Then

$$f(a+h) = f(a) + f'(a) \cdot h + \tfrac{1}{2!} f''(a) \cdot h^2 + \tfrac{1}{3!} f'''(a) \cdot h^3 + \cdots + \mathcal{O}(h^n) \tag{1.1}$$

And now the backward direction. Set $x = a - h$ in the general Taylor polynomial to obtain

$$f(a-h) = f(a) - f'(a) \cdot h + \tfrac{1}{2!} f''(a) \cdot h^2 - \tfrac{1}{3!} f'''(a) \cdot h^3 \pm \cdots + \mathcal{O}(h^n) \tag{1.2}$$

This idea can be easily extended to generate Taylor polynomials to use for approximations for higher order derivatives. Look at the Taylor polynomials resulting from setting $x = a + 2h$ and $x = a - 2h$. For example,

$$f(a+2h) = f(a) + 2f'(a) \cdot h + 2f''(a) \cdot h^2 + \tfrac{4}{3} f'''(a) \cdot h^3 + \cdots + \mathcal{O}(h^n),$$

and so forth.

1.4.2 Forward Difference Approximation

Stand at the point $x_0 = a$ and "look forward" to $x_1 = a + h$. Generate the Forward Difference Formula from (1.1).

To calculate the slope of the chord through $(a, f(a))$ and $(a + h, f(a + h))$, subtract $f(a)$ from both sides of (1.1) with $n = 2$, then divide by h to obtain:

$$\frac{f(a+h) - f(a)}{h} = f'(a) + \frac{\mathcal{O}(h^2)}{h}$$

Recall that $\mathcal{O}(h^2)/h = \mathcal{O}(h)$.

The Forward Difference Formula with error $\mathcal{O}(h)$ is

$$f'(a) = \frac{f(a+h) - f(a)}{h} + \mathcal{O}(h) \tag{1.3}$$

Compare (1.3) *with the limit definition of the derivative.*

The reason for using (1.1) rather than a simple slope formula is that we have the error term $\mathcal{O}(h)$ giving us an estimate of how good our approximation to the derivative is.

The error, $\mathcal{O}(h)$, that we have identified is called *truncation error*. Truncation error results from using a Taylor polynomial, a finite number of terms of the series. A deeper analysis of error would include *round-off error*, the error that results from using "floating point" arithmetic—an implementation that has a maximum, finite number of digits. For example, $1/3 + 1/3 + 1/3 = 1$, but $0.33 + 0.33 + 0.33 = 0.99 \neq 1$. There is competition between truncation error and round-off error making finding the right balance in a real problem non-trivial. For example, a detailed analysis gives the following theorem for error with the forward difference formula.

Theorem Let $f(x)$ have two continuous derivatives near $x = a$. Set

$$M_0 = \max_{a \le x \le a+h} |f(x)| \quad \text{and} \quad M_2 = \max_{a \le x \le a+h} \left| f''(x) \right|.$$

Then

$$\left| f'(a) - \frac{f(a+h) - f(a)}{h} \right| \le \underbrace{\frac{M_2}{2} \cdot h}_{\text{truncation error}} + \underbrace{2M_0 \varepsilon_\mu \cdot \frac{1}{h}}_{\text{round-off error}}$$

where ε_μ is determined by the floating point arithmetic machine used.

Observe the interplay: a smaller h makes truncation error smaller, but makes round-off error larger as shown in Fig. 1.3. Finding an optimal value is often not easy.

To study error more deeply than we will be doing enters the realm of Numerical Analysis courses.

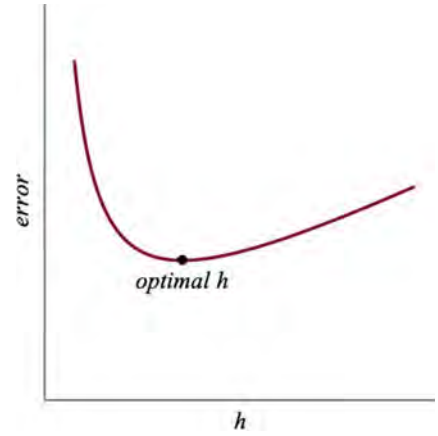


Fig. 1.3 The Error Graph

Forward Difference Calculations

1. Suppose $f(x) = 1 + x\, e^{\sin(x)}$. For $a = 0$ and $h = 0.1$, we have
$$f'(0) \approx \frac{1.1105 - 1.0000}{0.1} = 1.1050$$
2. Let $g(x) = \tan(\sin(x))$. For $a = 0.5$ and $h = 0.1$, we have
$$g'(0.5) \approx \frac{0.6334 - 0.5199}{0.1} = 1.1350$$
3. Suppose a function k is given by the data $P_0 = (1.000, 3.320)$ and $P_1 = (1.100, 3.682)$. Then
$$k'(1.00) \approx \frac{3.682 - 3.320}{1.100 - 1.000} = 3.620$$

Check the approximations above for f' and g' using algebraic differentiation to find exact values!

We can define a very simple Maple function to compute forward difference calculations for us.

A Maple Function to Calculate Forward Differences.

We'll choose the syntax to be

ForwardDiff(< *function name* >, < *x-value* >, < *step size* >).

```
> ForwardDiff := (f, a, h) -> (f(a+h) - f(a))/h :
```

```
> f := x -> x^2 - x - 1 :
  ForwardDiff(f, 2, 0.05);
  f'(2);
                          3.050000000
                                3
```

Extra for Maple Experts: Modify *ForwardDiff* to accept either a function or an expression as its first argument.

A Set of Parabolas Passing Through (1, 1) **and** (2, 2)**.**

There are an infinite number of parabolas passing through the two points (1, 1) and (2, 2). If we take $a = 1$ and $h = 1.0$, then the forward difference approximation for all these functions is the same

$$f'(1) \approx \frac{2-1}{1.0} = 1.0.$$

Look at what happens to the approximations to four different parabolas when $h = 0.5$ versus when $h = 1.0$ in Fig. 1.4.

The actual values of the four parabolas' derivatives at $a = 1.0$ are 2, 3, 4, and 5, respectively.

Create parabolas with positive concavity through the two points (1, 1) *and* (2, 2) *and do the same computations as above.*

1.4.3 Backward Difference Approximation

Now stand at the point $x_0 = a$ and "look backward" to $x_1 = a - h$. This generates the Backward Difference Formula.

To calculate the slope of the chord through $(a - h, f(a - h))$ and $(a, f(a))$, subtract $f(a)$ from both sides of (1.2) with $n = 2$, then divide by h to obtain the formula:

$$\frac{f(a-h) - f(a)}{h} = -f'(a) + \frac{\mathcal{O}(h^2)}{h}$$

As before, $\mathcal{O}(h^2)/h = \mathcal{O}(h)$.

The Backward Difference Formula with error $\mathcal{O}(h)$ is

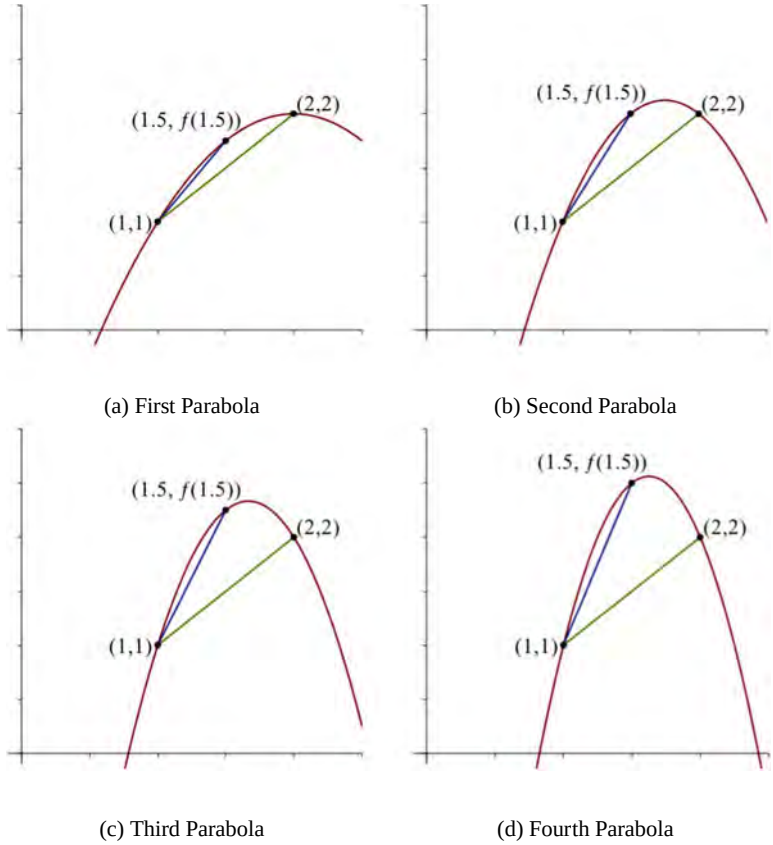
$$f'(a) = \frac{f(a) - f(a-h)}{h} + O(h) \tag{1.4}$$

Again the reason for using (1.2) is so that we have an estimate of how good our approximation to the derivative is. The error gets smaller as we make h smaller.

Backward Differences Examples

1. Again, suppose $f(x) = 1 + x\, e^{\sin(x)}$. For $a = 0.0$ and $h = 0.1$, we have

$$f'(0.0) \approx \frac{1.0000 - 0.910}{0.1} = 0.900$$



(a) First Parabola (b) Second Parabola

(c) Third Parabola (d) Fourth Parabola

Fig. 1.4 Four Parabolas Through (1, 1) and (2, 2) and Forward Differences

2. Let $g(x) = \tan(\sin(x))$ once more. For $a = 0.5$ and $h = 0.1$, we have

$$g'(0.5) \approx \frac{0.5199 - 0.4104}{0.1} = 1.0950$$

3. Let a function k be given by data: $P_0 = (1.000, 3.320)$ and $P_{-1} = (0.900, 2.970)$. Then

$$k'(1.000) \approx \frac{3.320 - 2.970}{1.000 - 0.900} = 3.500$$

Compare the answers above with those from algebraic differentiation!

We can also define a very simple Maple function to perform backward difference computations for us.

A Maple Function to Calculate Backward Differences.

We'll choose the syntax to be

$$BackwardDiff(< function\ name >, < x\text{-}value >, < step\ size >).$$

```
> BackwardDiff := (f, a, h) -> (f(a) - f(a - h))/h :
```

```
> f := x -> x^2 - x - 1 :
  BackwardDiff(f, 2, 0.05);
  f'(2);
                              2.950000000
                                    3
```

Extra for Maple Experts: Modify the function to accept expressions rather than functions.

Redux: A Set of Parabolas Passing Through $(1, 1)$ **and** $(2, 2)$**.**

Now consider the backward difference approximation to the parabolas passing through the two points $(1, 1)$ and $(2, 2)$. See Fig. 1.5. Take $a = 2$ and $h = 1.0$. Then

$$f'(2) \approx \frac{2-1}{1.0} = 1.0.$$

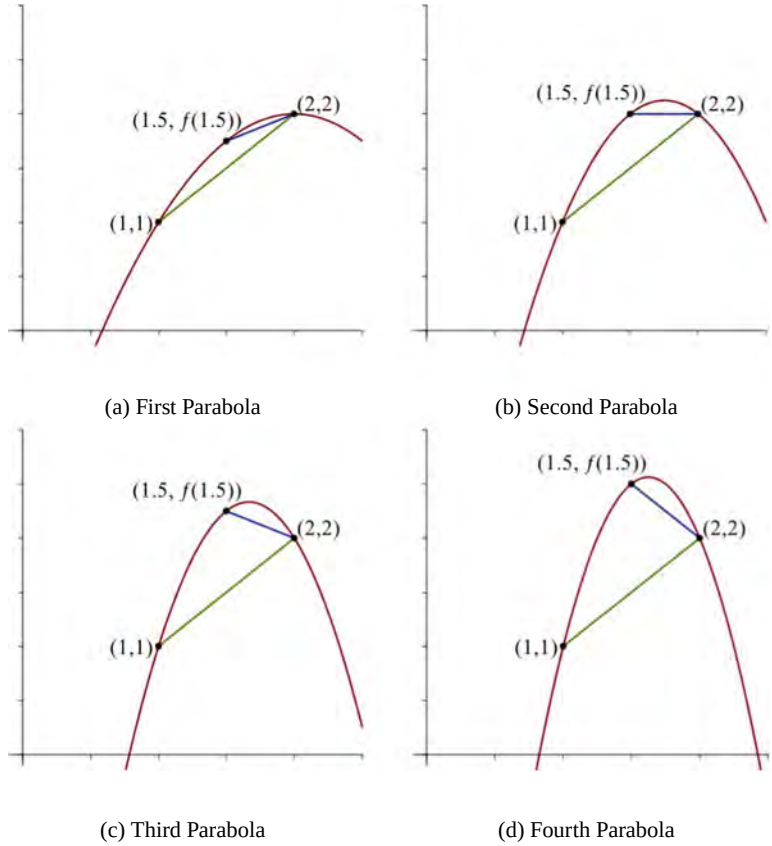
Consider what happens to the approximations to four different parabolas when $h = 0.5$ versus when $h = 1.0$ in Fig. 1.5. Now, the actual values of the four parabolas' derivatives at $a = 2.0$ are $0, -1, -2$, and -3, respectively. This time, $f'(2) \approx 1.0$ is a particularly bad estimate—geometry matters!

1.4.4 Centered Difference Approximation

The Centered Difference Formula comes from computing the slope of the chord that connects the "looking backward" point $a - h$ to the "looking forward" point $a + h$. The centered difference is also called the symmetric difference. Texas Instruments uses this formula with the default value $h = 0.001$ for the `nDeriv` function in their calculators.[6] Casio and HP calculators also use centered differences to calculate numerical derivatives.

Subtract $\mathcal{O}(h^3)$ versions of (1.1) and (1.2).

[6] See *TI-84 Plus and TI-84 Plus Silver Edition Guidebook*, pg. 40; a PDF is available from https://education.ti.com/-/media/6CC4C5AED5004F808892046AD33D4A35.



(a) First Parabola (b) Second Parabola

(c) Third Parabola (d) Fourth Parabola

Fig. 1.5 Four Parabolas Through (1, 1) and (2, 2) and Backward Differences

$$\begin{aligned} f(a+h) &= f(a) + f'(a)\cdot h + \tfrac{1}{2!}f''(a)\cdot h^2 + \mathcal{O}\left(h^3\right) \\ f(a-h) &= f(a) - f'(a)\cdot h + \tfrac{1}{2!}f''(a)\cdot h^2 + \mathcal{O}\left(h^3\right) \\ \hline f(a+h) - f(a-h) &= \qquad + 2f'(a)\cdot h + \qquad\qquad \mathcal{O}\left(h^3\right) \end{aligned}$$

Notice that the h^2 terms cancel! Remember that $O\left(h^3\right) - O\left(h^3\right) = O\left(h^3\right)$.

Solve for $f'(a)$ to obtain

$$f'(a) = \frac{f(a+h) - f(a-h)}{2h} + \frac{O(h^3)}{h}.$$

Remember that $\mathcal{O}(h^3)/h = \mathcal{O}(h^2)$. The Centered Difference Formula with error given by $\mathcal{O}(h^2)$ is

$$f'(a) = \frac{f(a+h) - f(a-h)}{2h} + O(h^2) \tag{1.5}$$

Since the h^2 terms canceled, the formula is more accurate than either forward or backward differences.

Centered Difference Examples

1. Once more, suppose $f(x) = 1 + x\, e^{\sin(x)}$. For $a = 0$ and $h = 0.1$, we have
$$f'(0.0) \approx (1.110 - 0.910)/0.2 = 1.000$$
2. Let $g(x) = \tan(\sin(x))$. For $a = 0.5$ and $h = 0.1$, we have
$$g'(0.5) \approx \frac{0.6334 - 0.4104}{2 \cdot 0.1} = 1.1150$$
3. Let a function k be given by data: $P_0 = (1.000, 3.320)$, $P_1 = (1.100, 3.682)$, and $P_{-1} = (0.900, 2.970)$. Then
$$k'(1.000) \approx \frac{3.682 - 2.970}{2 \cdot 0.1} = 3.560$$

Again compare the answers above with those from algebraic differentiation!

Next we'll define a very simple Maple function that computes centered difference approximations for us.

A Maple Function to Calculate Centered Differences.

We'll keep the pattern by choosing the syntax to be

CenteredDiff(< *function name* >, < *x-value* >, < *step size* >).

```
> CenteredDiff := (f, a, h) -> (f(a+h) - f(a-h))/(2*h):
```

```
> f := x -> x^2 - x - 1:
  CenteredDiff(f, 2, 0.05);
  f'(2);
                          3.000000000
                                  3
```

Extra for Maple Experts: Modify the function to accept an expression for f.

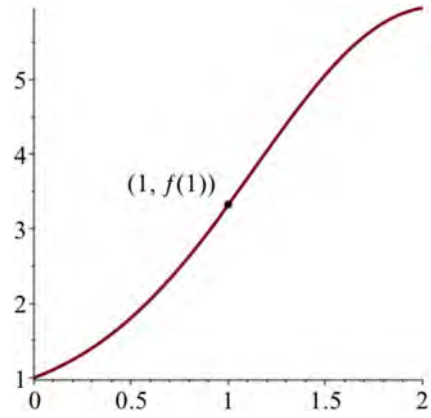


Fig. 1.6 Graph of $f(x) = 1 + x\, e^{\sin(x)}$

A Table of Differences From a Function

Let $f(x) = 1 + x\, e^{\sin(x)}$ and $a = 1.0$. Then

$$f'(1.0) = e^{\sin(1)}\,(1 + \cos(1)) \approx 3.573157593$$

Figure 1.6 shows a graph of f. Table 1.2 shows the results of using our three formulas, forward difference, backward difference, and centered difference to estimate $f'(1)$ with varying h. *Verify the computations!*

Table 1.2 Differences From a Function

h	FD(h)	*error*	BD(h)	*error*	CD(h)	*error*
2^{-1}	3.494890	0.078268	3.024408	0.548750	3.259649	0.313509
2^{-2}	3.636316	0.063158	3.347764	0.225394	3.492040	0.081118
2^{-3}	3.628464	0.055306	3.476944	0.096214	3.552704	0.020454
2^{-4}	3.606368	0.033210	3.529696	0.043462	3.568032	0.005126
2^{-5}	3.591104	0.017946	3.552640	0.020518	3.571872	0.001286
2^{-6}	3.582464	0.009306	3.563200	0.009958	3.572832	0.000326
2^{-7}	3.577600	0.004442	3.568256	0.004902	3.572928	0.000230

Table 1.3 The Function's Data

x_i	−1.60	−1.20	−0.80	−0.40	0.00	0.40	0.80	1.20	1.60
y_i	−0.29	0.56	0.81	0.65	0.30	−0.06	−0.21	0.04	0.89

Derivatives from Data

Many times a function is given by a table of values. In real applications, having data is much, much more likely than a formula. We can still approximate derivatives by calculating the differences with the given values, but our step size h may be fixed by the data.

Estimate the derivatives of a function given the data below in Table 1.3 where $h = 0.4$.

Figure 1.7 shows a plot of Table 1.3's data.

Use the data with all three formulas: forward, backward, and centered difference (Table 1.4). The symbol $\bowtie$ marks entries in the table where a derivative cannot be calculated with the respective formula. *Why can't they be calculated?*

The actual derivatives listed in Table 1.5 were computed with the function's formula—a cubic polynomial. *Can you determine the formula?*

We close this section with a caveat. Be careful that the function is defined at $x = a$ and that there is a derivative at $x = a$—the centered difference formula doesn't use $f(a)$. For example, the absolute value function has no derivative at $x = 0$, but the centered difference formula for $f(x) = |x|$ at $a = 0$ always simplifies to 0. The function $g(x) = 1/x$ is undefined

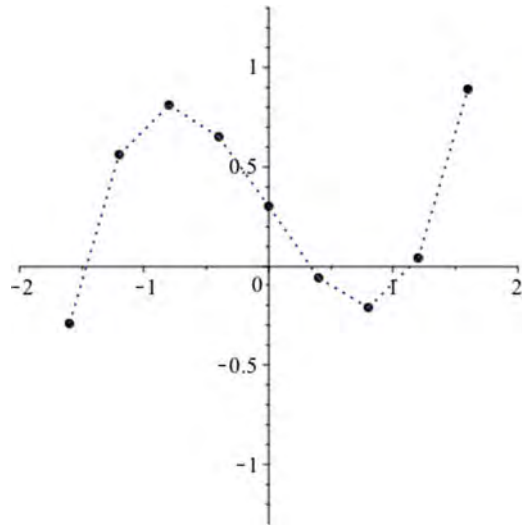


Fig. 1.7 Graph of Table 1.3's Data

Table 1.4 Estimates of the Derivatives

Forward Differences

x_i	−1.60	−1.20	−0.80	−0.40	0.00	0.40	0.80	1.20	1.60
FD(x_i)	2.12	0.625	−0.375	−0.90	−0.90	−0.375	0.625	2.12	⋈

Backward Differences

x_i	−1.60	−1.20	−0.80	−0.40	0.00	0.40	0.80	1.20	1.60
BD(x_i)	⋈	2.12	0.625	−0.375	−0.90	−0.90	−0.375	0.625	2.12

Centered Differences

x_i	−1.60	−1.20	−0.80	−0.40	0.00	0.40	0.80	1.20	1.60
CD(x_i)	⋈	1.374	0.125	−0.637	−0.90	−0.6375	0.125	1.374	⋈

Table 1.5 Actual Values of the Derivatives

x_i	−1.60	−1.20	−0.80	−0.40	−0.005	0.40	0.80	1.20	1.60
y_i'	3.057	1.293	0.033	−0.723	−0.975	−0.723	0.033	1.293	3.057

Table 1.6 Difference method error

Forward difference	$\mathcal{O}(h)$
Backward difference	$\mathcal{O}(h)$
Centered difference	$\mathcal{O}(h^2)$

at $x = 0$, but the centered difference of g at $x = 0$ for any h is $CD(h) = 1/h^2$. Always check for the existence of the function and its derivative!

1.4.5 Comparing the Error Terms

Both the forward difference (1.3) and backward difference (1.4) have error term $\mathcal{O}(h)$ as shown in Table 1.6. This error term indicates that if we divide h by 10, we can expect to divide the error by 10; i.e., we get one more decimal place of accuracy. Contrast this with the centered difference (1.5) that has error term $\mathcal{O}(h^2)$. Here we expect dividing h by 10 will divide the error by $10^2 = 100$, two more decimal places of accuracy—quite a change!

But take care—as noted above, the centered difference formula will give $y'(0) = 0$ when $y(x) = |x|$ for any $h > 0$ even though y is not differentiable at $x = 0$. *Try this on a graphing calculator!*

An even more dramatic example comes from using the centered difference formula to find $f'(0)$ for the function $f(x) = x\sin(1/x)$. All the centered differences at $a = 0$ are equal to 0, yet the derivative of f goes to infinity as a goes to 0. Look at several graphs to see what's happening.

The moral: good analysis of a function is required before using any of the approximation methods.

1.5 Higher Derivatives

We can use the formulas in equations (1.3), (1.4), and (1.5) to derive corresponding formulas for higher order derivatives. The process is recursive, so we must be very careful with the variables. For a preview of the technique, rewrite (1.3) with g as

$$\frac{d}{dx}g(x) = \frac{g(x+h) - g(x)}{h}.$$

Replace g with f':

$$\begin{aligned}\frac{d}{dx}f'(x) &= f''(x)\\ &= \frac{f'(x+h) - f'(x)}{h}.\end{aligned}$$

Now replace each f' term using (1.3) to generate a formula approximating $f''(x)$. The full set of steps follows.

Forward Difference Formula for the Second Derivative
Begin by applying (1.3) to $f''(x)$ to have

$$f''(x) = \frac{f'(x+h) - f'(x)}{h} + \mathcal{O}(h).$$

Replace the f' terms using (1.3).

$$f''(x) = \frac{\dfrac{f((x+h)+h) - f(x+h)}{h} - \dfrac{f(x+h) - f(x)}{h}}{h} + \mathcal{O}(h)$$

Simplify to obtain a forward difference formula for the second derivative, $f''(x)$.

$$f''(x) = \frac{f(x+2h) - 2f(x+h) + f(x)}{h^2} + \mathcal{O}(h). \tag{1.6}$$

Backward Difference Formula for the Second Derivative

Repeat the process above instead using (1.4) to have

$$f''(x) = \frac{f'(x) - f'(x-h)}{h} + \mathcal{O}(h).$$

Replace the f' terms using (1.4).

$$f''(x) = \frac{\dfrac{f(x) - f(x-h)}{h} - \dfrac{f(x-h) - f((x-h)-h)}{h}}{h} + \mathcal{O}(h).$$

Then simplify to obtain a backward difference formula for $f''(x)$.

$$f''(x) = \frac{f(x) - 2f(x-h) + f(x-2h)}{h^2} + \mathcal{O}(h). \tag{1.7}$$

Centered Difference Formula for the Second Derivative

One more time, repeat the process but use (1.5) in the calculations.

$$\begin{aligned} f''(x) &= \frac{f'(x+h) - f'(x-h)}{2h} + \mathcal{O}(h) \\ [1ex] &= \frac{\dfrac{f((x+h)+h) - f((x+h)-h)}{2h} - \dfrac{f((x-h)+h) - f((x-h)-h)}{2h}}{2h} + \mathcal{O}(h^2). \end{aligned}$$

Now simplify to yield a centered difference formula for $f''(x)$.

$$f''(x) = \frac{f(x+2h) - 2f(x) + f(x-2h)}{4h^2} + \mathcal{O}(h^2). \tag{1.8}$$

The formula we just found is essentially equivalent to a formula directly derived using the Taylor series methods we began with:

$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} + \mathcal{O}(h^2).$$

Explain how the two formulas are equivalent!

The process we've been using can be repeated to create formulas for 3rd and higher derivatives. *Be very careful with the* '$\pm h$'s!

There are many methods for numerically calculating derivatives that are more intricate than those we have considered. Start investigating these techniques by looking at *interpolation and numerical differentiation* and *Newton's differences*; see, e.g., [5] and [6].

1.6 Next Steps

There are a number of directions for further study of numerical derivatives. For a starting point, choose:

- *Lagrange Interpolation,* a technique to efficiently fit a polynomial to data. The resulting polynomial can easily be differentiated providing approximate derivatives for the data.
- *Richardson's Extrapolation,* a technique to generate higher accuracy estimations by combining values from lower order formulas.

The field of *numerical analysis* considers these and many other techniques and problems of computation.

Exercises

1. Is the following statement true? Carefully justify your answer.
   If f has a derivative at $x = a$, then $f(x) = \mathcal{O}(x)$ as $x \to a$.
2. Let $f(x) = \cos(x)$ and set $a = 0$. Use the listed methods to find $f'(a)$ with

   a. a forward difference with step size $h = 0.1$ and $h = 0.01$,
   b. a backward difference with step size $h = 0.1$ and $h = 0.01$,
   c. a centered difference with step size $h = 0.1$ and $h = 0.01$;

3. Let $f(x) = \sin(x^2 + \pi/2)$ and set $a = \sqrt{\pi/2}$. Use the listed methods to find

   a. $f'(a)$ with

      i a forward difference with step size $h = 0.1$ and $h = 0.01$,
      ii a backward difference with step size $h = 0.1$ and $h = 0.01$,
      iii a centered difference with step size $h = 0.1$ and $h = 0.01$;

   b. $f''(a)$ with

      i a forward difference with step size $h = 0.1$ and $h = 0.01$,
      ii a backward difference with step size $h = 0.1$ and $h = 0.01$,
      iii a centered difference with step size $h = 0.1$ and $h = 0.01$.

4. Derive the Uniform Form that bounds R_n from each of the other forms in Table 1.1 (sect. 1.1.1).
5. Let $f(x) = \tan(\sin(x))$.

a. Graph f over the interval $[-\pi, \pi]$.
b. Try to find a Taylor polynomial that is a good approximation to f over the whole interval. Can it be done? What is happening?
c. *Extra for Experts:* Plot $z = |f(x + iy)|$ on the square $[-\pi, \pi] \times [-\pi, \pi]$ to illustrate the problem.

6. Show that the Centered Difference Formula (1.5) is the average of the Forward and Backward Difference Formulas (1.3) and (1.4).
7. Explain why the Centered Difference Formula (1.5) is $\mathcal{O}(h^2)$ rather than $\mathcal{O}(h)$.
8. Use $\mathcal{O}(h^4)$ versions of equations (1.1) and (1.2) to find a centered difference approximation for $f''(a)$.
9. Investigate the ratio of error in the function's difference chart (Sect. 1.4.4) as h is successively divided by 2 for

a. forward differences
b. backward differences
c. centered differences

10. Examine the ratios of error to h in Table 1.3, the data differences chart, (Sect. 1.4.5) for

a. forward differences
b. backward differences
c. centered differences

11. Use Maple to produce Table 1.2, Differences from a Function. The table can be created using a *for* loop and the three difference functions we defined.
12. Rewrite the three Maple difference functions to accept either a function or an *expression* as input.
13. A calculator's `nDeriv` function (syntax: `nDeriv`(*expr*, *var*, *val*, <*step*>)) with a default step of $h = 0.001$ can be implemented in Maple as[7]

```
> nDeriv := proc(y, x, a, h := 0.001)
      (eval(y, x = a + h) - eval(y, x = a - h))/(2h);
  end proc :
```

Change the Maple function *nDeriv* to use

a. forward differences,
b. backward differences.

Apply *nDeriv* to random polynomials to compare the accuracy.

[7] Maple's $eval(y, x = z)$ evaluates the expression y after substituting z for x.

14. Define the *unit step function*,

$$U(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}.$$

Let h be any positive value. Compute

a. the forward difference formula to estimate $H'(0)$,
b. the backward difference formula to estimate $H'(0)$,
c. the centered difference formula to estimate $H'(0)$.
d. What is $H'(0)$?

15. Derive the 5-*point difference formula* for $f'(a)$ by combining Taylor expansions to $\mathcal{O}(h^5)$ for $f(a)$, $f(a \pm h)$, and $f(a \pm 2h)$.
16. Develop forward, backward, and centered difference formulas for $f'''(x)$ using the methods of Sect. 1.5.
17. Compare the coefficients in forward difference formulas for $f'(x)$, $f''(x)$, $f'''(x)$, and $f''''(x)$. Can you identify a pattern?
18. Collect the temperatures (e.g., with a CBL) in a classroom from 8:00 am to 5:00 pm.

a. Estimate the rate of change of temperatures during the day.
b. Compare plots of the rates given by forward, backward, and centered differences.

19. Collect cost v. time data as a gas pump fills an automobile.

a. Graph the data.
b. Estimate the rate of flow of gas given by forward, backward, and centered differences.

20. G.W. Bickley's paper "Formulæ for Numerical Differentiation" [7, pp. 19-27], presents tables of coefficients for derivative approximation formulas. Investigate the paper and report on the formulas.
21.

a. Find Taylor expansions for sin and cos to $\mathcal{O}(x^6)$. Estimate $\cos(1.0)$.
b. Since $\frac{d}{dx}\sin(x) = \cos(x)$, we can estimate cos with the derivative of sin. Use your expansion of sin and $h = 0.05$ to approximate $\cos(1.0)$ with

i. forward differences
ii. backward differences
iii. centered differences

Discuss the errors.

c. Research how most calculators compute sine and cosine values. *Hint: It's not with Taylor polynomials!*

22. Investigate the connection between repeated synthetic division and Taylor polynomials.
23. Use Maple to create a table of the standard trigonometric functions with their Taylor polynomials pointing out the patterns that appear in the polynomials' coefficients for each function.
24. Investigate Taylor polynomials for two functions that are close to each other. Define

$$f_1(x) = \tan(1) \cdot \sin(x) \qquad \text{and} \qquad f_2(x) = \tan(\sin(x))$$

a. Graph the two functions together for $-2\pi \leq x \leq 2\pi$.
b. Use Maple to compute the Taylor polynomials for f_1 and f_2 up to degree 7.
c. Define the two Maple functions:

```
> FrmS := n →
      plot([f1(x), mtaylor(f1(x), x, n + 1)], x = −2 · Pi..2 · Pi, −4..4) :
  FrmTS := n →
      plot([f2(x), mtaylor(f2(x), x, n + 1)], x = −2 · Pi..2 · Pi, −4..4) :
```

d. Now use Maple's *Explore* command to graph the functions with their Taylor polynomials using the degree as a control.

```
> Explore(FrmS(n), n = 1..15);
```

and

```
> Explore(FrmTS(n), n = 1..15);
```

e. Describe your observations.

25. Investigate using numerical derivatives for edge detection in image processing.

Appendices

1.7 Taylor's Statement of the Theorem

In 1712, Brook Taylor wrote a letter to John Machin containing his theorem, but without a proof. Then the theorem appeared with a proof in Taylor's book *Methodus Incrementorum Directa et Inversa*[8] (1715) as Corollary II to Proposition VII. The proposition is a restatement of "Newton's [interpolation] Formula."

[8] See Ian Bruce's annotated translation of Taylor's *Methodus Incrementorum Directa et Inversa*.

Colin Maclaurin introduced the method, *order of contact—undetermined coefficients,* in Sect. 751 of *A Treatise of Fluxions* (1742) that we use now to present Taylor's theorem in elementary calculus classes.

Corollary 1.1 (Taylor's Corollary II *[pg 23]*) *If for the evanescent increments, the fluxions that are proportional to them are written, the quantities* $\ddot{v}$, $\dot{v}$, v, $\dot{v}$, $\ddot{v}$, &c. *being now made all equal to the time* z *uniformly flows to become* $z + v$, *then* x *will become*

$$x + \dot{x}\,\frac{v}{1\,\dot{z}} + \ddot{x}\,\frac{v^2}{1\,.\,2\,\dot{z}^2} + \dddot{x}\,\frac{v^3}{1\,.\,2\,.\,3\,\dot{z}^3} + \&c.$$

Note that, in just under the thirty years between Taylor's book being published and Maclaurin's, scientific writing had begun moving from Latin to native languages, here English.

There was some controversy in giving Taylor credit for the theorem discussed in several papers; some thought Bernoulli deserved more credit, others Sterling or Hermann. A nice summary can be read in [8]. Persson et al [9] give a summary of the remainder term along with several proofs of Taylor's Theorem.

1.8 Centered Difference Coefficients Chart

In 1988, Bengt Fornberg published an algorithm[9] to compute the coefficients of difference formulas for derivatives of any order or accuracy. Fornberg's algorithm is only 17 lines and is easy to implement. Coding the algorithm makes a very good student project. We'll reproduce Table 1 [10, pg 703] of his paper that shows the coefficients for centered difference formulas for up to the fourth derivative as Table 1.7.

Each row of Fornberg's table gives the coefficients of the centered difference formula approximating a desired derivative to a specified order. For example, using the ninth row of Table 1.7 gives the *third* derivative's centered difference approximation with *second-order* or $\mathcal{O}(h^2)$ accuracy as

$$f^{(3)}(a) \approx \frac{-\frac{1}{2}f(a-2h) + 1f(a-h) + 0f(a) - 1f(a+h) + \frac{1}{2}f(a+2h)}{h^3} + \mathcal{O}(h^2).$$

The coefficients in the table are not difficult to verify using expansions like (1.1) and (1.2) (Sect. 1.4). Follow a procedure similar to our development of the centered difference formula. Project 3.6 (Sect. 3.6) studies Fornberg's algorithm.

[9] Bengt Fornberg, "Generation of Finite Difference Formulas on Arbitrarily Spaced Grids," *Math of Comp* **51** No. 184 (Oct., 1988), pp 699–706.

Table 1.7 Centered Finite Difference Formula Coefficients

Derivative	$\mathcal{O}(h^a)$	$x-4h$	$x-3h$	$x-2h$	$x-h$	x	$x+h$	$x+2h$	$x+3h$	$x+4h$
1	2				$-\frac{1}{2}$	0	$\frac{1}{2}$			
	4			$\frac{1}{12}$	$-\frac{2}{3}$	0	$\frac{2}{3}$	$-\frac{1}{12}$		
	6		$-\frac{1}{60}$	$\frac{3}{20}$	$-\frac{3}{4}$	0	$\frac{3}{4}$	$-\frac{3}{20}$	$\frac{1}{60}$	
	8	$\frac{1}{280}$	$-\frac{4}{105}$	$\frac{1}{5}$	$-\frac{4}{5}$	0	$\frac{4}{5}$	$-\frac{1}{5}$	$\frac{4}{105}$	$-\frac{1}{280}$
2	2				1	-2	1			
	4			$-\frac{1}{12}$	$\frac{4}{3}$	$-\frac{5}{2}$	$\frac{4}{3}$	$-\frac{1}{12}$		
	6		$\frac{1}{90}$	$-\frac{3}{20}$	$\frac{3}{2}$	$-\frac{49}{18}$	$\frac{3}{2}$	$-\frac{3}{20}$	$\frac{1}{90}$	
	8	$-\frac{1}{560}$	$\frac{8}{315}$	$-\frac{1}{5}$	$\frac{8}{5}$	$-\frac{205}{72}$	$\frac{8}{5}$	$-\frac{1}{5}$	$\frac{8}{315}$	$-\frac{1}{560}$
3	2			$-\frac{1}{2}$	1	0	-1	$\frac{1}{2}$		
	4		$\frac{1}{8}$	-1	$\frac{13}{8}$	0	$-\frac{13}{8}$	1	$-\frac{1}{8}$	
	6	$-\frac{7}{240}$	$\frac{3}{10}$	$-\frac{169}{120}$	$\frac{61}{30}$	0	$-\frac{61}{30}$	$\frac{169}{120}$	$-\frac{3}{10}$	$\frac{7}{240}$
4	2			1	-4	6	-4	1		
	4		$-\frac{1}{6}$	2	$-\frac{13}{2}$	$\frac{28}{3}$	$-\frac{13}{2}$	2	$-\frac{1}{6}$	
	6	$\frac{7}{240}$	$-\frac{2}{5}$	$\frac{169}{60}$	$-\frac{122}{15}$	$\frac{91}{8}$	$-\frac{122}{15}$	$\frac{169}{60}$	$-\frac{2}{5}$	$\frac{7}{240}$

1.9 Links and Others

There are large numbers of websites with information about numerical derivatives. A Google search returns over 66 million pages in less than half a second. A Google Scholar search returns over 4.5 million pages in less than a tenth of a second. A Bing search returns over 66 million pages in less than half a second.

- Web searches for "numerical derivative":
  - Google
  - Google Scholar
  - Bing
- Science Direct's Numerical Differentiation Topic Page:
  - www.sciencedirect.com/topics/mathematics/numerical-differentiation
- Encyclopedia of Mathematics from Springer Verlag entry "Differentiation, numerical"
  - encyclopediaofmath.org/wiki/Differentiation,_numerical

- SpringerLink's "Minimal numerical differentiation formulas"
  - link.springer.com/article/10.1007/s00211-018-0973-3
- Online Engineering Notes "Interpolation and approximation numerical differentiation"
  - onlineengineeringnotes.com/2021/05/26/interpolation-and-approximation-numerical-differentiation-and-regression/
- MacTutor's biographies of Brook Taylor:
  - mathshistory.st-andrews.ac.uk/Biographies/Taylor/
  - mathshistory.st-andrews.ac.uk/Strick/taylor.pdf

References

1. Uhl, J., Davis, W., Porta, H.: Calculus & Mathematica. Addison-Wesley (1994)
2. Hairer, E., Wanner, G.: Analysis by its History. Springer-Verlag (1996)
3. Bauldry, W.C.: Introduction to Real Analysis: An Educational Approach. John Wiley & Sons (2011)
4. Boyer, C.B.: The History of the Calculus and Its Conceptual Development. Courier Corporation (1959)
5. Burden, R., Faires, J.D.: Numerical Analysis, 10th edn. Brooks/Cole Cengage Learning (2016)
6. Kahaner, D., Moler, C., Nash, S.: Numerical Methods and Software (1989)
7. Bickley, W.G.: Formulae for numerical differentiation. The Mathematical Gazette **25**(263), 19–27 (1941)
8. Gibson, G.A.: Taylor's Theorem and Bernoulli's Theorem : A Historical Note. Proceedings of the Edinburgh Mathematical Society **39**, 25–33 (1920)
9. Persson, L.E., Rafeiro, H., Wall, P.: Historical synopsis of the Taylor remainder. Note di Matematica **37**(1), 1–22 (2017)
10. Fornberg, B.: Generation of finite difference formulas on arbitrarily spaced grids. Mathematics of Computation **51**(184), 699–706 (1988)

Check for updates

Numerical Integration

2

What is *numerical integration*? Numerical integration or *(numerical)* quadrature is the calculation of the value of a definite integral using numerical formulas, not the fundamental theorem with antiderivatives. The Greeks studied quadrature: given a figure, construct a square that has the same area. The two most famous are Hippocrates of Chios' *Quadrature of the Lune* (c. 450 BC) and Archimedes' *Quadrature of the Parabola* (c. 250 BC). Archimedes used the method of exhaustion—a precursor to calculus—invented by Eudoxus of Cnidas (Fig. 2.1). Archimedes also discovered the quadrature of the surface of a spheroid.[1] *Squaring the circle* is one of the classical quadrature problems of constructing a square with the area of a given circle; it was shown to be impossible by Carl von Lindemann in 1882.[2]

2.1 Methods of Elementary Calculus

The first methods we learn for numerically integrating a function come directly from what we've seen in calculus texts as motivation for understanding what a definite integral can be interpreted as geometrically. Start by constructing simple geometric figures approximating the desired area. The easiest shape is a rectangle with the simple formula $Area = height \times width$. We'll look at the formulas, their error bounds, and ways to improve our area estimates in the following sections.

[1] Archimedes, "On the Sphere and Cylinder".

[2] Lindemann was the first to prove that https://proofwiki.org/wiki/Pi_is_Transcendental π is transcendental. *Explore* Lindemann's Theorem.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

W. C. Bauldry, *Computational Calculus*, Synthesis Lectures on Mathematics & Statistics,
https://doi.org/10.1007/978-3-031-29658-1_2

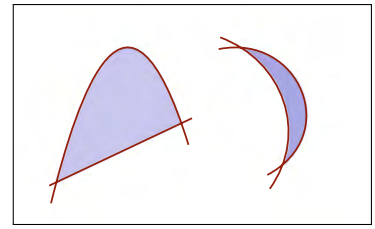


Fig. 2.1 Archimedes' Parabola and Hippocrates' Lune

2.1.1 Rectangle Methods

The *rectangle methods* approximate the height of a changing function with a constant; i.e., we replace the function over a small interval with a rectangle enclosing approximately the same amount of area. Using the flat line of a rectangle for the function's height is called a "zero-order approximation" from $y = a_0x^0$; using a line would be a "first-order" approximation from $y = a_0x^0 + a_1x^1$, and so forth. This technique is motivated by the definition of a definite integral as the limit[3] of Riemann sums $\sum_k f(c_k)\Delta x_k$. The different methods come from choosing a different place to measure the height of the function. A *left endpoint approximation* uses the height at the left end of each rectangle or *panel*. A *right endpoint approximation* uses the height at the right end of each panel. And a *midpoint* approximation uses the height at the midpoint of each panel.

Suppose we wish to find the area under f shown in Fig. 2.2.
Figure 2.3 shows the three basic rectangle methods for approximating the area defined by the function f.

In general, none of the rectangles will give the exact area—the goodness of approximation depends on the function and the width of the rectangle. Although, the *Mean Value Theorem for Integrals* tells us that a continuous function does have an x value whose rectangle would give the exact area—but this is an "existence theorem" which doesn't give any insight into finding that special x value. Often, carefully subdividing into narrower and narrower rectangles can get us as close as we like to the actual value.

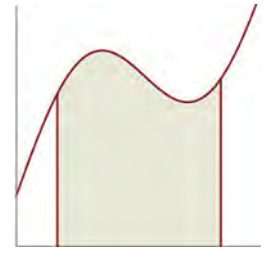
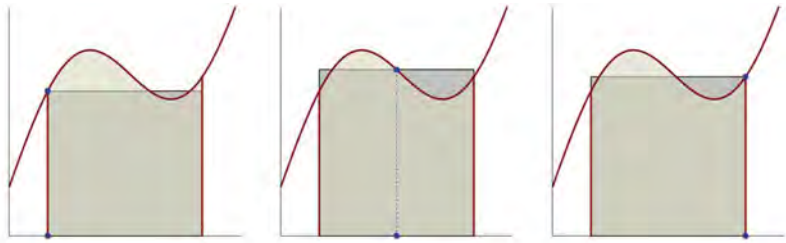


Fig. 2.2 The desired area under the curve f

[3] There is a great deal of subtlety in the idea of a "limit of Riemann sums." Analyzing this concept is begun in advanced calculus studies.



(a) Left Endpoint Height (b) Midpoint Height (c) Right Endpoint Height

Fig. 2.3 The rectangle methods

Left Endpoint Approximation

Break the interval $[a, b]$ into n rectangles $[x_{i-1}, x_i]$ for i from 1 to n where $x_0 = a$ and $x_n = b$. The width of each rectangle is $\Delta x_i = (x_i - x_{i-1})$. Measure the height of each rectangle at its left endpoint as in Fig. 2.3a. The ith rectangle's area is $f(x_{i-1})\Delta x_i$. So the *left endpoint sum* approximating the area under f over $[a, b]$ is

$$L_n = \sum_{i=1}^{n} f(x_{i-1})\Delta x_i.$$

Note that taking equal width rectangles makes $\Delta x_i = (b - a)/n$ which simplifies the formula to

$$L_n^* = \frac{b-a}{n} \sum_{i=1}^{n} f(x_{i-1}).$$

Since $x_i = a + i\,\Delta x$, then

$$L_n^* = \frac{b-a}{n} \sum_{i=1}^{n} f(a + (i-1)\Delta x) = \frac{b-a}{n} \sum_{k=0}^{n-1} f(a + k\Delta x). \tag{2.1}$$

There are situations, however, where it's advantageous to let the rectangles have different widths.

Since L_n^* is an estimate of the actual value of the definite integral, we need to know about the potential error to make it a useful approximation—an estimate without an error analysis is of little to no use! Once again, Taylor polynomials lead us to the answer. Consider one rectangle: for $x \in [x_{i-1}, x_i]$,

$$f(x) = f(x_{i-1}) + f'(c_i)\Delta x$$

where c_i is some value $c_i \in [x_{i-1}, x_i]$.

Now integrate to see

$$\begin{aligned}\int_{x_{i-1}}^{x_i} f(x)\,dx &= \int_{x_{i-1}}^{x_i} \big(f(x_{i-1}) + f'(c_i)\Delta x\big)dx \\ &= f(x_{i-1}) \int_{x_{i-1}}^{x_i} dx + f'(c_i)\Delta x \int_{x_{i-1}}^{x_i} dx \\ &= f(x_{i-1})\,\Delta x + \tfrac{1}{2} f'(c_i)\Delta x^2.\end{aligned}$$

Summing all the rectangles to approximate the area from a to b gives

$$\int_a^b f(x)\,dx = L_n^* + \tfrac{1}{2} \sum_{i=1}^{n} f'(c_i)\Delta x^2$$

Since $\Delta x = (b-a)/n$, we have

$$\int_a^b f(x)\,dx = L_n^* + \frac{(b-a)^2}{2} \sum_{i=1}^{n} f'(c_i)\frac{1}{n^2} \tag{2.2}$$

Require f to have a derivative that is bounded on $[a, b]$ so there is a value M_1 such that $|f'(c_i)| \le M_1$, and recognize that $\sum_{k=1}^{n} M_1/n^2 = M_1/n$ to let us write the error term using Big-$\mathcal{O}$.

$$\begin{aligned}\int_a^b f(x)\,dx &= L_n^* + \frac{(b-a)^2}{2} \cdot M_1 \cdot \frac{1}{n} \\ &= L_n^* + \mathcal{O}(1/n).\end{aligned}$$

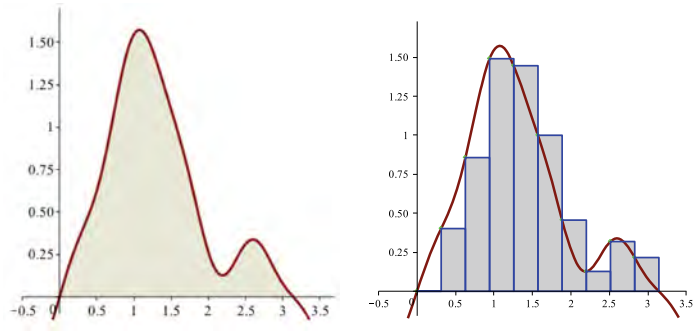
The error for L_n^* is $\mathcal{O}(\Delta x) = \mathcal{O}(1/n)$.

An error of $\mathcal{O}(1/n)$ indicates that if we double the number of rectangles, we can expect to halve the error. For instance, if the error in approximating the area with 20 left endpoint rectangles was less than 0.01, then the error should be less than 0.005 for a sum of 40 left endpoint rectangles.

However, (2.2) gives us even more information for a good error estimate. Let $|f'(x)| \le M_1$ on $[a, b]$, then the error ε_n with L_n^* is bounded by

$$\varepsilon_n \le \frac{(b-a)^2}{2} \cdot M_1 \cdot \frac{1}{n}. \tag{2.3}$$

Examine how each part, the length of the interval, a bound for the derivative, and the number of rectangles, effects the error estimate. Given values for a, b, and M_1, we can determine an n that will guarantee our error is less than a specified tolerance.



(a) The Area (b) The Left Endpoint Approximation

Fig. 2.4 The left endpoint sum

A Left Endpoint Sum

Let

$$f(x) = \sin(x) + \frac{1}{2}\sin(2x) - \frac{1}{4}\sin(4x) + \frac{1}{16}\sin(8x).$$

Approximate the area under f over the interval $[0, \pi]$ with 10 equal width rectangles in a left endpoint sum.

First, a graph of f is shown in Fig. 2.4a.

We will use the formula for L_{10}^*,

$$L_{10}^* = \sum_{k=1}^{10} f(x_{k-1})\,\Delta x.$$

Set $\Delta x = (\pi - 0)/10$, the length of the interval divided by 10 steps.

What are the x_k? We have 10 equal steps of Δx from $x_0 = a = 0.0$ to $x_{10} = b = \pi$, thus

$$x_k = a + k\Delta x = \frac{k\pi}{10}.$$

Then, to five places,

$$\begin{aligned} L_{10}^* &= 0.31416 \sum_{k=1}^{10} f(0.31416(k-1)) \\ &= 1.98352 \end{aligned}$$

Now, for our error estimate, Fig. 2.5 shows a graph of $|f'(x)|$ on $[0, \pi]$. Maple produced this plot nicely with the command

```
> plot({2.3, |f'(x)|}, x = 0..Pi, 0..2.5)
```

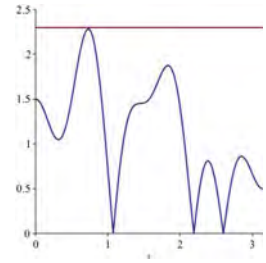


Fig. 2.5 Graph of $|f'(x)|$ and $y = 2.3$

From the graph, we can see that $M_1 = 2.3$ is a suitable bound for f' on $[0, \pi]$. Therefore, our upper bound for the error to five places from using 10 left endpoint rectangles is

$$\begin{aligned}\varepsilon_{10} &\leq \frac{(\pi - 0)^2}{2} \cdot 2.3 \cdot \frac{1}{10} \\ &\leq 1.13500\end{aligned}$$

That's fairly large! *How could the bound be improved?* Since 2.3 is a large overestimate nearly everywhere on the interval, the real error is much smaller. This type of estimate for the derivative leads to quite conservative choices for the number of rectangles. The exact value of the integral is 2.

Let's reverse the direction and ask how many rectangles are necessary.

Determining the Number of Rectangles

Suppose again that

$$f(x) = \sin(x) + \frac{1}{2}\sin(2x) - \frac{1}{4}\sin(4x) + \frac{1}{16}\sin(8x).$$

How many equal width rectangles in a left endpoint sum are needed to approximate the area under f over the interval $[0, \pi]$ with an error less than 10^{-2}?

We have already determined that $M_1 = 2.3$ for $x \in [0, \pi]$. Put these values in the error inequality to have the error ε_n satisfies

$$\varepsilon_n \leq \frac{(\pi - 0)^2}{2} \cdot 2.3 \cdot \frac{1}{n} \leq 10^{-2}.$$

Solving for n gives $1135.0045 < n$, so any number of rectangles greater than or equal to 1136 must work.

Note that this is a very conservative error estimate—smaller values of n will likely work, but 1136 is guaranteed to work.

Right Endpoint Approximation

In right endpoint approximation, we switch to using the right endpoint of the interval $[x_{i-1}, x_i]$ as the place to measure the height of f. The ith rectangle's area is now $f(x_i)\Delta x_i$. So the *right endpoint sum* approximating the area under f over $[a, b]$ is

$$R_n = \sum_{i=1}^{n} f(x_i)\Delta x_i.$$

Again note that taking equal width rectangles makes $\Delta x_i = (b-a)/n$ which simplifies the formula to

$$R_n^* = \frac{b-a}{n} \sum_{i=1}^{n} f(x_i). \tag{2.4}$$

We had determined that for n equal-width rectangles, $x_i = a + i\,\Delta x$, therefore

$$R_n^* = \frac{b-a}{n} \sum_{i=1}^{n} f(a + i\,\Delta x).$$

The error analysis for right endpoint sum is essentially the same as for the left endpoint sum. Require f to have a derivative that is bounded on $[a, b]$, and find a value M_1 such that $|f'(c_i)| \le M_1$. Once again recognize that $\sum_{k=1}^{n} M_1/n^2 = M_1/n$. Then the error term using Big-$\mathcal{O}$ is

$$\begin{aligned} \int_a^b f(x)\,dx &= R_n^* + \frac{(b-a)^2}{2} \cdot M_1 \cdot \frac{1}{n} \\ &= R_n^* + \mathcal{O}(1/n). \end{aligned}$$

That is, the error for R_n^* is $\mathcal{O}(\Delta x) = \mathcal{O}(1/n)$, or, specifically,

$$\varepsilon_n \le \frac{(b-a)^2}{2} \cdot M_1 \cdot \frac{1}{n}. \tag{2.5}$$

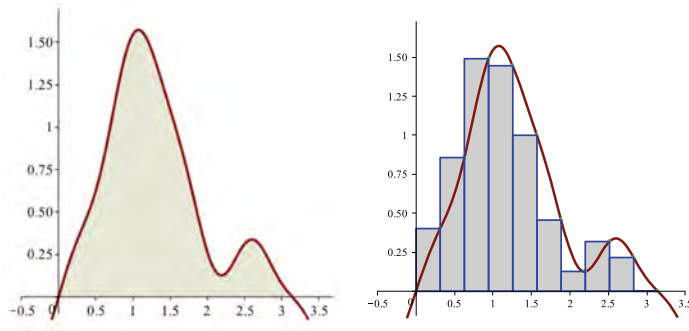
This error bound shouldn't be a surprise—comparing Figs. 2.4b and 2.6b, we see that only two rectangles' areas don't appear in both sums, $[x_0, x_1] \cdot \Delta x$ and $[x_9, x_{10}] \cdot \Delta x$. That is, all but the first and last rectangles appear in both the left endpoint and the right endpoint sums.

Let's revisit the left endpoint sum example, but use right endpoint sums.

A Right Endpoint Sum

Once again, set

$$f(x) = \sin(x) + \frac{1}{2}\sin(2x) - \frac{1}{4}\sin(4x) + \frac{1}{16}\sin(8x).$$



(a) The Area (b) The Right Endpoint Approximation

Fig. 2.6 The right endpoint sum

Approximate the area under f over the interval $[0, \pi]$ with 10 equal width rectangles in a right endpoint sum.

The formula for R_{10}^* is

$$R_{10}^* = \sum_{k=1}^{10} f(x_k)\,\Delta x,$$

where $\Delta x = (\pi - 0)/10$, the length of the interval divided by 10 steps.

Since there are 10 equal steps of Δx from $x_0 = 0.0$ to $x_{10} = \pi$, then

$$x_k = x_0 + k\Delta x = \frac{k\pi}{10}.$$

Then, to five places,

$$\begin{aligned} R_{10}^* &= 0.31416 \sum_{k=1}^{10} f(0.31416\,k) \\ &= 1.98352 \end{aligned}$$

Why does the right endpoint sum equal the left endpoint sum for this function!

The computation of the number of rectangles for a right endpoint sum is the same as it was for left endpoint sum, so we'll leave it to the exercises.

Midpoint Approximation

Measure the height of each rectangle at its midpoint as in Fig. 2.3b. The ith rectangle's area is $f(x_{c_k})\Delta x_i$ where $c_k = (x_k - x_{k-1})/2$. The *midpoint sum* approximating the area under f over $[a, b]$ is

$$C_n = \sum_{i=1}^{n} f\left(\frac{x_i - x_{i-1}}{2}\right) \Delta x_i.$$

Note that taking equal width rectangles makes $\Delta x_i = (b-a)/n$ again simplifying the formula to

$$C_n^* = \frac{b-a}{n} \sum_{i=1}^{n} f\left(\frac{x_i - x_{i-1}}{2}\right). \tag{2.6}$$

Recognize that the midpoints occur at $c_k = a + (k - 1/2)\Delta x$. Then

$$C_n^* = \frac{b-a}{n} \sum_{k=1}^{n} f\left(a + \left(k - \frac{1}{2}\right)\Delta x\right).$$

Surprisingly, measuring the rectangle's height at the midpoint of each interval can give a more accurate estimate of the integral that the left or right endpoint.

Taylor polynomials will again lead us to an error bound. Consider the error for one rectangle $x \in [x_{i-1}, x_i]$. The midpoint formula gives this area as $f(c_i)(x_i - x_{i-1})$. So the error ε_i fort he ith rectangle is

$$\begin{aligned}\varepsilon_i &= \int_{x_{i-1}}^{x_i} f(x)\,dx - f(c_i)\cdot(x_i - x_{i-1}) \\ &= \int_{x_{i-1}}^{x_i} f(x) - f(c_i)\,dx.\end{aligned}$$

From Taylor's theorem, $f(x) - f(c_i) = f'(c_i)(x - c_i) + f''(\xi)(x - c_i)^2$ for some $\xi \in [x_{i-1}, x_i]$.So

$$\begin{aligned}\varepsilon_i &= \int_{x_{i-1}}^{x_i} f'(c_i)(x - c_i) + f''(\xi)(x - c_i)^2\,dx \\ &= f'(c_i)\int_{x_{i-1}}^{x_i} (x - c_i)\,dx + f''(\xi)\int_{x_{i-1}}^{x_i} (x - c_i)^2\,dx \\ &= f'(c_i)\cdot 0 \quad + f''(\xi)\cdot\frac{(x_i - x_{i-1})^3}{24}.\end{aligned}$$

Let M_2 be a real number such that

$$\max_{x\in[a,b]} \left|f''(x)\right| \le M_2.$$

Now, $(x_i - x_{i-1})^3 = ((b-a)/n)^3$ for each rectangle, so adding all the errors gives

$$\sum_{i=1}^{n} |\varepsilon_i| \le \sum_{i=1}^{n} M_2 \cdot \frac{(b-a)^3}{24\,n^3}$$

$$\varepsilon \le \frac{(b-a)^3}{24} \cdot M_2 \cdot \frac{1}{n^2} = \mathcal{O}(h^2). \tag{2.7}$$

Explain how the midpoint sum has $\mathcal{O}(n^{-2})$, while the left and right endpoint sums have $\mathcal{O}(n^{-1})$!

A Midpoint Sum

Let's use our familiar test function

$$f(x) = \sin(x) + \frac{1}{2}\sin(2x) - \frac{1}{4}\sin(4x) + \frac{1}{16}\sin(8x).$$

Approximate the area under f over the interval $[0, \pi]$ with 10 equal width rectangles in a midpoint sum. See Fig. 2.7.

The formula for C_{10}^* is

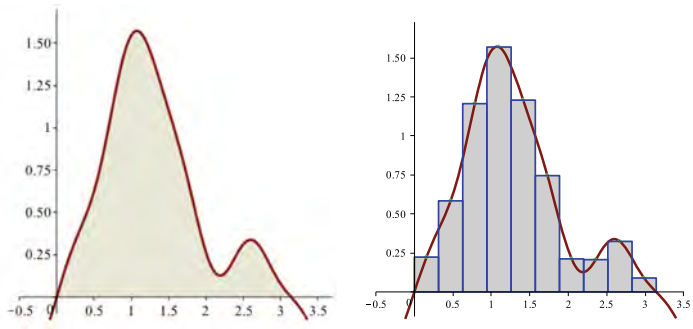
$$C_{10}^* = \sum_{k=1}^{10} f(c_k)\,\Delta x,$$

where $\Delta x = (\pi - 0)/10$, the length of the interval divided by 10 steps.

Since there are 10 equal steps of Δx from $x_0 = 0.0$ to $x_{10} = \pi$, the midpoint between x_{i-1} and x_i is

$$c_i = x_{i-1} + \frac{1}{2}\Delta x;$$

that is,



(a) The Area (b) The Midpoint Approximation

Fig. 2.7 The midpoint sum

$$
\begin{aligned}
c_1 &= x_0 + \frac{1}{2}\Delta x \\
c_2 &= x_1 + \frac{1}{2}\Delta x = x_0 + \frac{3}{2}\Delta x \\
c_3 &= x_2 + \frac{1}{2}\Delta x = x_0 + \frac{5}{2}\Delta x \\
&\vdots
\end{aligned}
$$

So

$$c_k = x_0 + \left(k - \frac{1}{2}\right)\Delta x = 0 + \left(k - \frac{1}{2}\right) \cdot \frac{\pi}{10}.$$

Then, to five places,

$$
\begin{aligned}
C_{10}^{*} &= 0.31416 \sum_{k=1}^{10} f(0.31416\,k - 0.157080) \\
&= 2.00825
\end{aligned}
$$

The error estimate requires a value bounding the second derivative of f's absolute value. Use Maple to produce a graph

```
> plot([9.75, |f''(x)|], x = 0..π);
```

showing that $M_2 = 9.75$ is a convenient bound for $|f''(x)|$. The graph appears in Fig. 2.8. The error bound for a midpoint sum gives (to 6 decimals)

$$\varepsilon_{10} \le \frac{(\pi - 0)^3}{24} \cdot 9.75 \cdot \frac{1}{10^2} = 0.125963.$$

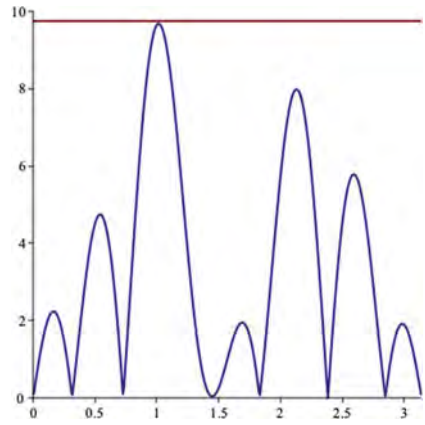


Fig. 2.8 Bounding the absolute value of f''

That is much better than what we had for the left or right endpoint sums!

Compare this computation to the exact value of the integral, 2.

Let's reverse direction and ask how many midpoint rectangles are needed to guarantee the error is less than 10^{-2}.

Determining the Number of Panels Needed for a Given Error

Use the error formula with ε_n:

$$\varepsilon_n \leq \frac{(\pi - 0)^3}{24} \cdot 9.75 \cdot \frac{1}{n^2} \leq 10^{-2}.$$

Solving for n gives $35.491266 < n$. So any midpoint sum with 36 or more rectangles must have an error less 10^{-2}.

Remember that our error formula is *very* conservative. Figure 2.8 shows that the value $M_2 = 9.75$ is very far above the curve nearly everywhere over the interval $[0, \pi]$. If we were to break the interval into several subintervals, we could have better bounds with different numbers of rectangles for different areas—this idea leads to what is called *adaptive quadrature*.

Summary of the Rectangle Methods

It's time to collect the methods and error formulas for comparison. Table 2.1 presents a summary of the rectangle methods formulas and error bounds.

All the rectangle methods are actually special versions of a general Riemann sum

$$\mathcal{R}_n = \sum_{k=1}^{n} f(c_k)\,\Delta x_k \quad \text{where} c_k \in [x_{k-1}, x_k]$$

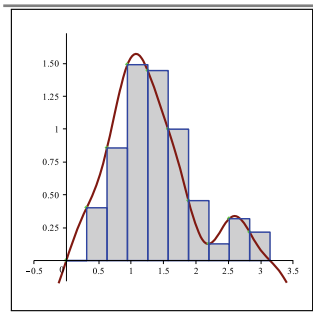
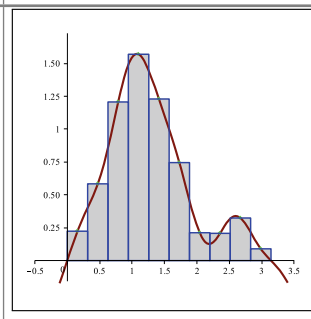
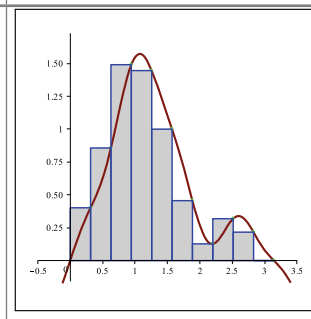
that we use to define the definite integral. The typical definition

$$\int_a^b f(t)\,dt = \lim_{\|\mathcal{P}_n\| \to 0} \sum_{k=1}^{n} f(c_k)\,\Delta x_k \quad \text{where} \mathcal{P}_n = \{x_k \,:\, k = 0..n\}$$

hides a great deal of subtlety that is studied in Advanced Calculus or Introduction to Real Analysis courses.

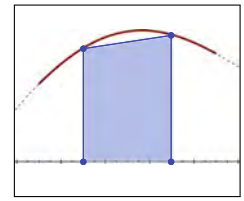
Now it's time to leave rectangles and go to the next higher order approximation of $f(x)$.

Table 2.1 A Summary of the rectangle methods

Left endpoint sum	Midpoint sum	Right endpoint sum
		
$A_n \approx \sum_{k=1}^{n} f(x_{k-1})\,\Delta x_k$	$A_n \approx \sum_{k=1}^{n} f(m_k)\,\Delta x_k$	$A_n \approx \sum_{k=1}^{n} f(x_k)\,\Delta x_k$
	$m_k = \frac{1}{2}(x_{k-1} + x_k)$	
$\varepsilon_n \le \dfrac{(b-a)^2}{2} \cdot M_1 \cdot \dfrac{1}{n}$	$\varepsilon_n \le \dfrac{(b-a)^3}{24} \cdot M_2 \cdot \dfrac{1}{n^2}$	$\varepsilon_n \le \dfrac{(b-a)^2}{2} \cdot M_1 \cdot \dfrac{1}{n}$

where $M_i = \max \left| f^{(i)}(x) \right|$;

2.1.2 Trapezoid Sums



Instead of the rectangle or zero degree approximations to the function, use linear or first degree approximations. We now have trapezoids rather than rectangles. The area of the trapezoid over the interval $[x_{k-1}, x_k]$ is given by

$$A_T = \tfrac{1}{2}[f(x_{k-1}) + f(x_k)]\,\Delta x_k$$

Summing the trapezoids gives an approximation for the integral.

$$\int_a^b f(x)\,dx \approx \sum_{k=1}^{n} \tfrac{1}{2}[f(x_{k-1}) + f(x_k)]\,\Delta x_k$$

The formula is often written as

$$T_n \approx \left[f(x_0) + \left(2 \sum_{k=1}^{n-1} f(x_k) \right) + f(x_n) \right] \cdot \frac{\Delta x_k}{2}. \tag{2.8}$$

Note that, while the midpoint estimate measured the function's height at the average x, the trapezoid approximation averages the function's height at the endpoints; trapezoids average the y values, while midpoints average the x values.[4]

It will be simpler to use a theorem from *interpolation theory* to derive an error bound for a trapezoid sum than to use Taylor series. The interpolation theorem[5] states that when f has a continuous second derivative, the line $L(x)$ through the two points $(x_{i-1}, f(x_{i-1}))$ and $(x_i, f(x_i))$ satisfies

$$f(x) - L(x) = \frac{1}{2} \cdot f''(c_x) \cdot (x - x_{i-1})(x - x_i)$$

for all $x \in [x_{i-1}, x_i]$. Intuitively, this theorem says that a parabola with the right concavity sitting on the line $L(x)$ captures the function; that is, the function is between the parabola and the line L.

Integrate both sides of the interpolation error equation from $x = x_{i-1}$ to x_i and take absolute values.

$$\left| \int_{x_{i-1}}^{x_i} f(x)\,dx - \int_{x_{i-1}}^{x_i} L(x)\,dx \right| = \left| \int_{x_{i-1}}^{x_i} \frac{1}{2} \cdot f''(c_i) \cdot (x - x_{i-1})(x - x_i)\,dx \right|$$

Let M_2 be greater than $|f''(x)|$ over $[a, b]$, then

$$\begin{aligned} \left| \int_{x_{i-1}}^{x_i} f(x)\,dx - \int_{x_{i-1}}^{x_i} L(x)\,dx \right| &\le \frac{1}{2} \cdot M_2 \cdot \left| \int_{x_{i-1}}^{x_i} (x - x_{i-1})(x - x_i)\,dx \right| \\ &\le \frac{1}{2} \cdot M_2 \cdot \frac{(x_i - x_{i-1})^3}{6} \\ &\le \frac{1}{12} \cdot M_2 \cdot h^3 \end{aligned}$$

Since $L(x)$ is the line through $(x_{i-1}, f(x_{i-1}))$ and $(x_i, f(x_i))$, the integral of $L(x)$ is just the area of the trapezoid that interpolation line defines. Thus

$$\left| \int_{x_{i-1}}^{x_i} f(x)\,dx - \frac{h}{2} \cdot (f(x_{i-1}) + f(x_i)) \right| \le \frac{1}{12} \cdot M_2 \cdot h^3.$$

Sum over all n trapezoids.

[4] In a similar fashion, changing the focus from x to y led Henri Lebesgue to develop *measure theory* and the *Lebesgue integral*.

[5] See, e.g., [1, Sect. 3.6].

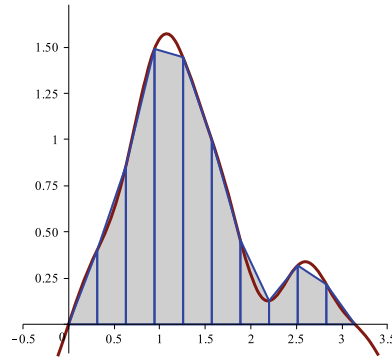


Fig. 2.9 Trapezoid rule approximation

$$\left| \int_a^b f(x)\,dx - \frac{h}{2} \cdot \sum_{i=1}^{n} (f(x_{i-1}) + f(x_i)) \right| \leq \frac{1}{12} \cdot M_2 \cdot h^3 \sum_{i=1}^{n} 1$$

Where has the Triangle inequality[6] *come into play in the last two inequalities?*

Replace h with $(b-a)/n$, and we have developed the error bound for the trapezoid sum as

$$\varepsilon_n \leq \frac{(b-a)^3}{12} \cdot M_2 \cdot \frac{1}{n^2}. \tag{2.9}$$

Looking at the formulas shows that a trapezoid sum is the average of the left and right endpoint sums. Carefully consider why the error bound for the trapezoid sum is $\mathcal{O}(h^2)$ while the bounds for left and right endpoint sums are only $\mathcal{O}(h)$. What could it be about averaging left- and right-endpoint sums that cancels some error?

A Trapezoid Sum

Let $f(x) = \sin(x) + \frac{1}{2}\sin(2x) - \frac{1}{4}\sin(4x) + \frac{1}{16}\sin(8x)$ over $[0, \pi]$.

We'll take $n = 10$, and then the equipartition nodes are $x_k = k\pi/10$ for $k = 0..10$.

With this partition,

$$\Delta x = \frac{\pi}{10} \approx 0.314159$$

Then Fig. 2.9 shows that

$$T_{10} = \left[f(0) + \left(2 \sum_{k=1}^{9} f(\tfrac{k}{10}\pi) \right) + f(\pi) \right] \frac{\Delta x}{2}$$

which gives

$$T_{10} = 1.983524.$$

[6] The triangle inequality, $|a+b| \leq |a| + |b|$, is one of the most important in mathematical analysis.

Recall that Fig. 2.8 showed $M_2 = 9.75$. The computed maximum error from the trapezoid error formula is

$$\begin{aligned}\varepsilon_{T_{10}} &\leq \frac{(\pi - 0)^3}{12} \cdot 9.75 \cdot \frac{1}{10^2} \\ &\leq 0.251925.\end{aligned}$$

From the exact value we computed earlier, the actual absolute error is 0.016476. The trapezoid sum worked quite well with only 10 panels. We again see that the actual error is much smaller than what the error bound formula provides since the estimate for M_2 is *very* conservative.

An Example of Roundoff Error

Before we continue with integration methods, we'll take a brief detour into roundoff error. Recall that when we encountered truncation error for numerical derivatives, we observed that there is balance between the smallness of the step size h and roundoff from numerical calculations: as step size became smaller, roundoff became more important. The analogous situation in numerical integration is the smallness of Δx balancing roundoff from numerical computation. Consider the following example of the integral of $\int_0^1 \sin(\pi x)\, dx$. Analytically, this integral is equal to $2/\pi$. Watch what happens as we use more and more trapezoids.

Balancing Roundoff Error and Δx

Execute the steps below in Maple.[7]

```
> f := x → sin(evalf(Pi) · x) :
  T := n → 1/(2n) · sum(evalf(f((k−1)/n)) + evalf(f(k/n)), k = 1..n) :
  v := evalf(int(f(x), x = 0..1), 12);
                        v := 0.636619772284
```

[7] You may need to adjust the precision in a different CAS to see this behaviour. TI calculators use two hidden "guard digits" and hide this problem. However, on calculators, it's still easy to see that $(\sqrt{3})^2 \neq 3$ due to roundoff error. For more information on calculators, look at TI's https://education.ti.com/~/media/5EF5EF4CAA204719B0FAB17076D8D16C (PDF download link); also see [2] and the *Case Study* in [3].

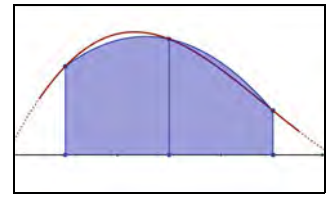
$$> Matrix(10, 2, [seq([n = 2^i, T_n = |v - T(2^i)|], i = 4..13)]);$$

$$\begin{bmatrix} n = 16 & T_n = 0.0020466238 \\ n = 32 & T_n = 0.0005114093 \\ n = 64 & T_n = 0.0001278368 \\ n = 128 & T_n = 0.0000319563 \\ n = 256 & T_n = 7.9883 \times 10^{-6} \\ n = 512 & T_n = 1.9968 \times 10^{-6} \\ n = 1024 & T_n = 1.7948 \times 10^{-6} \\ n = 2048 & T_n = 0.0000260097 \\ n = 4096 & T_n = 0.0000803157 \\ n = 8192 & T_n = 0.0002969267 \end{bmatrix}$$

Look carefully at the values of T_n to see that the error begins to increase even though Δx is getting smaller!

Just as we saw with step size in our formulas for numerically differentiating a function, there is an optimal number n of trapezoids (or rectangles, etc.) to use to minimize error. Finding the optimal number of panels can be quite challenging and involves different competing factors: the number of rectangles which determines the number of computations versus the amount of precision or the number of decimals needed. For a very large problem, the amount of time required to perform computations can be an issue. These topics are studied in *numerical analysis*.

2.1.3 Simpson's Rule



Moving from the trapezoid's order 1 or linear approximation to order 2 or quadratic approximation to $f(x)$ brings us to Simpson's Rule.[8] Replace the trapezoid's line through the endpoints of each interval with a parabola. However, determining a unique parabola requires three data points. The easiest way to have 3 data points is to

[8] As is the case with many named mathematical results, Simpson wasn't the first to develop Simpson's rule. Simpson published the result in 1743, acknowledging that he had learned it from Isaac Newton. However, what we call Newton's Method was actually developed by Simpson, so there's some "fairness" to the name switch. Johannes Kepler derived the rule in 1615, a form of the rule was used by Bonaventura Cavalieri in 1639, and James Gregory reported it in 1668. See Goldstine's *A History of Numerical Analysis from the 16th through the 19th Century* [4] for the whole story.

take the panels in pairs: instead of the area base $[x_i, x_{i+1}]$, use $[x_i, x_{i+1}, x_{i+2}]$. So we require an even number of panels. The area under the parabola is easy to compute as $A_S = \frac{1}{3}\left[f(x_i) + 4f(x_{i+1}) + f(x_{i+2})\right]\Delta x$,
although the algebra is somewhat intense. We'll let Maple do that so we can concentrate on the concept.

A Lagrange interpolation formula will give us the unique parabola through the three points $(a, f(a))$, $(b, f(b))$, and $(c, f(c))$.

$$> p := t \to f(a)\cdot\frac{t-b}{a-b}\cdot\frac{t-c}{a-c} + f(b)\cdot\frac{t-a}{b-a}\cdot\frac{t-c}{b-c} + f(c)\cdot\frac{t-a}{c-a}\cdot\frac{t-b}{c-b}:$$

A quick check:

$$> p(a), p(b), p(c);$$
$$f(a), f(b), f(c)$$

Substitute $a = x_{k-1}, b = x_{k-1} + h$, and $c = x_{k-1} + 2h$ in $p(t)$, then integrate the resulting quadratic function from $t = x_{k-1}$ to $t = x_{k-1} + 2h$.

$$> s := subs([a = x_{k-1}, b = x_{k-1} + h, c = x_{k-1} + 2\cdot h], p(t)):$$
$$int(s, t = x_{k-1}..x_{k-1} + 2\cdot h):$$
$$sp := factor(expand(\%));$$
$$sp := \frac{h\left(f(x_{k-1}) + 4f(x_{k-1}+h) + f(x_{k-1}+2h)\right)}{3}$$

For even $N = 2n$, this construction gives the $2n$-panel approximation to the integral

$$\int_a^b f(x)\,dx \approx S_{2n} = \left[\sum_{k=1}^{n} f(x_{2k-2}) + 4f(x_{2k-1}) + f(x_{2k})\right]\cdot\frac{\Delta x}{3},$$

which is most often written as

$$S_{2n} = \left[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots + 4f(x_{2n-1}) + f(x_{2n})\right]\cdot\frac{\Delta x}{3}. \tag{2.10}$$

Simpson's Rule Integrates a Third Degree Polynomial Exactly

A surprising aspect of Simpson's rule is that there is no error when applying it to third degree polynomials even though it uses a second degree approximation to the function. For a test, Maple gave

$$p := x \to 42x^3 + 93x^2 + 33x + 79$$

as a random polynomial and $[-3, 2]$ as a random interval. Using only 2 panels (one parabola) to estimate the integral gives

$$\int_{-3}^{2} p(x)\,dx = 715 \qquad \text{and} \qquad S_2 = \frac{5}{6}\cdot\big(f(-3) + 4f(-1/2) + f(2)\big) = 715.$$

The exact value!

Exercise 9 (pg. 66) uses Maple to show algebraically why this must be true. However, looking at the error bound will easily explain why this surprise happens.

Simpson's Rule Error Bound

When f has a bounded fourth derivative, the basic error for Simpson's rule S_2 on $[x_{2k-2}, x_{2k}]$ is bounded by

$$\varepsilon_2 \le \frac{h^5}{90}\cdot M_4$$

where $M_4 \ge \max\left|f^{(4)}(x)\right|$. We can derive this result using the same methods that we found the bound for the trapezoid sum error. However, we would use the interpolation error formula for a fourth degree polynomial. The derivation is somewhat complex so we will leave it another course. See [1, pg. 74] or [5, pg. 280] for a proof.

To extend the error bound to $[a, b]$ from $[x_{2k-2}, x_{2k}]$ proceed as before: use ε_2 to find ε_N where N must be chosen to be an even integer number of panels with $a = x_0 < x_1 < \cdots < x_{2n} = b$.

Add the errors for all the panels remembering that we use them in pairs.

$$\left|\int_a^b f(x)\,dx - S_N\right| = \varepsilon_N \le \sum_{k=1}^{N/2} \frac{h^5}{90}\cdot M_4$$

Then

$$\varepsilon_N \le \frac{h^5}{90}\cdot M_4 \cdot \frac{N}{2}.$$

Now $h = (b-a)/N$, so

$$\varepsilon_N \le \frac{(b-a)^5}{180}\cdot M_4 \cdot \frac{1}{N^4}. \tag{2.11}$$

The reason for Simpson's rule integrating third degree polynomials is now easy to see: the error expression uses M_4. Since the fourth derivative of any cubic polynomial is zero, the error must be zero. Whence Simpson's rule has to integrate any third degree polynomial exactly!

For a fascinating method of analyzing the error in Simpson's rule, read Louis Talman's "Simpson's Rule Is Exact for Quintics" [6] (*but be careful to note what the title really means*). See Exercise 9 (pg. 66) for a direct computational proof that Simpson's rule must give the exact value for a cubic polynomial.

Let's revisit our familiar function so as to compare Simpson's rule results with our previous methods.

A Simpson Sample

Let $f(x) = \sin(x) + \frac{1}{2}\sin(2x) - \frac{1}{4}\sin(4x) + \frac{1}{16}\sin(8x)$ over $[0, \pi]$.

```
> S10 := ApproximateInt(f(x), x = 0..π, partition = 10, method = simpson) :
  evalf(S10);
```

$$S_{10} = 2.000006784$$

with absolute error of $6.78 \cdot 10^{-6}$.

Even just looking at the graphs of the methods indicates Simpson's rule is quite good in comparison. According to Philip Davis, "Abramowitz used to say—somewhat in jest—that 95% of all practical work in numerical analysis boiled down to applications of Simpson's rule and linear interpolation." [7, pg. 57]

2.1.4 A Maple Comparison

Let's consider the difficult integral[9]

$$\int_1^2 \frac{10^{-4}}{\left(x - \frac{1}{2}\pi\right)^2 + 10^{-8}}\, dx.$$

This integrand has a very sharp peak at $\pi/2$. *Graph the integrand!*
The exact value of the integral (using the https://openstax.org/books/calculus-volume-1/pages/5-3-the-fundamental-theorem-of-calculus *Fundamental Theorem of Calculus*) is

$$\arctan(5 \cdot 10^3 \cdot (4 - \pi)) - \arctan(5 \cdot 10^3 \cdot (2 - \pi)) \approx 3.1411844701381$$

The results from using our basic methods in Maple are collected in Table 2.2.

Looking at the values in the Table shows that to achieve a relative error below 1% requires approximately $n \geq 6000$.

The error bound for Simpson's rule, the best of the methods, uses M_4. Using a little calculus on $f^{(4)}(x)$ leads us to $M_4 = 4.1 \cdot 10^{29}$, a huge value! The error bound formula then says that to achieve an error less than 10^{-3} requires $4 \cdot 10^7$ panels! That sharp peak of the function overwhelms the analysis. Our best choice would be to break the interval $[1, 2]$ into several subintervals and deal with each separately. This thought leads to *adaptive quadrature*.

[9] A "test integral" from Berntsen, Espelid, & Sørevik [8].

Table 2.2 Maple's Evaluations

$n = 50$		$n = 500$		$n = 5000$	
left endpoint	0.0497133	*left endpoint*	0.541336	*left endpoint*	3.42282
right endpoint	0.0497180	*right endpoint*	0.541336	*right endpoint*	3.42282
midpoint	3.1210200	*midpoint*	4.052010	*midpoint*	2.88243
trapezoid	0.0497157	*trapezoid*	0.541336	*trapezoid*	3.42282
Simpson	2.0972500	*Simpson*	2.881790	*Simpson*	3.06256

Table 2.3 Numerical Integration Methods

Method	"Height"	Error Bound[a]
Left end point	$f(x_i)$	$\frac{1}{2}(b-a)^2 \cdot M_1 \cdot \frac{1}{n} = \mathcal{O}(h)$
Right end point	$f(x_{i+1})$	$\frac{1}{2}(b-a)^2 \cdot M_1 \cdot \frac{1}{n} = \mathcal{O}(h)$
Trapezoid Rule	$\frac{f(x_i)+f(x_{i+1})}{2}$	$\frac{1}{12}(b-a)^3 \cdot M_2 \cdot \frac{1}{n^2} = \mathcal{O}(h^2)$
Midpoint	$f\left(\frac{x_i+x_{i+1}}{2}\right)$	$\frac{1}{24}(b-a)^3 \cdot M_2 \cdot \frac{1}{n^2} = \mathcal{O}(h^2)$
Simpson's Rule	$\frac{f(x_i)+4f(x_{i+1})+f(x_{i+2})}{3}$	$\frac{1}{180}(b-a)^5 \cdot M_4 \cdot \frac{1}{n^4} = \mathcal{O}(h^4)$

[a] An approximation without an error bound has little to no value

2.1.5 A Chart of the Methods

Let's add Simpson's rule to the chart of collected the numerical integration methods, Table 2.3, for comparisons. As usual, we'll use $M_i \geq \max \left| f^{(i)}(x) \right|$ and $h = 1/n$ in the table. Notice that the methods can be classified as being *one point*, *two point*, or *three point* with error bounds $\mathcal{O}(h)$, $\mathcal{O}(h^2)$, or $\mathcal{O}(h^4)$, respectively, based on how the "height" is calculated.

These integration methods are all instances of *Newton-Cotes formulas*. We could go further: cubic polynomials give *Simpson's* $3/8$ *rule* (Simpson's rule above is often called *Simpson's* $1/3$ *rule*), fourth degree yields *Boole's rule*, and so forth.

2.2 The Basic Integration Methods and Maple

The *ApproximateIntTutor* program in the *Calculus1* subpackage of the *Student* package is very nice for quick comparisons and gaining intuition on the different methods. Load the function by executing *with*(*Student*[*Calculus1*]). Let's start *ApproximateIntTutor* with the function and range we've used for the initial methods:

$$f(x) = \sin(x) + \frac{1}{2}\sin(2x) - \frac{1}{4}\sin(4x) + \frac{1}{16}\sin(8x) \quad \text{over} \quad [0, \pi].$$

```
> with(Student[Calculus1]) :
```

```
> f := x → sin(x) + 1/2 sin(2x) − 1/4 sin(4x) + 1/16 sin(8x) :
```

Now execute *ApproximateIntTutor* to generate the window shown in Fig. 2.10.

```
> ApproximateIntTutor(f(x), x = 0..Pi) :
```

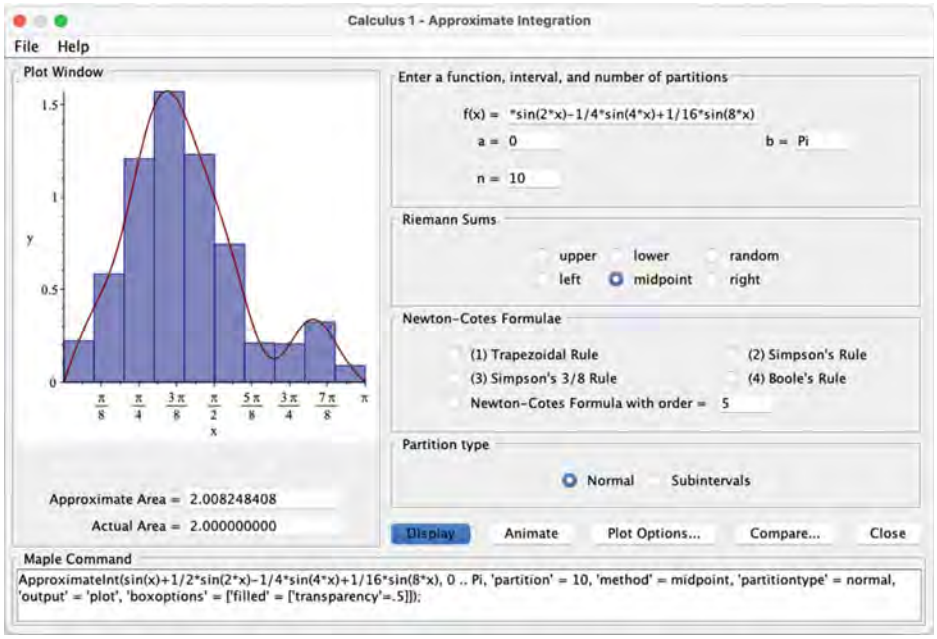

Calculus 1 - Approximate Integration
File Help
Plot Window
Enter a function, interval, and number of partitions
f(x) = "sin(2*x)-1/4*sin(4*x)+1/16*sin(8*x)
a = 0
b = Pi
n = 10
Riemann Sums
upper lower random
left midpoint right
Newton-Cotes Formulae
(1) Trapezoidal Rule
(2) Simpson's Rule
(3) Simpson's 3/8 Rule
(4) Boole's Rule
Newton-Cotes Formula with order = 5
Partition type
Normal Subintervals
Approximate Area = 2.008248408
Actual Area = 2.000000000
Display Animate Plot Options... Compare... Close
Maple Command
ApproximateInt(sin(x)+1/2*sin(2*x)-1/4*sin(4*x)+1/16*sin(8*x), 0 .. Pi, 'partition' = 10, 'method' = midpoint, 'partitiontype' = normal, 'output' = 'plot', 'boxoptions' = ['filled' = ['transparency'=.5]]);

Fig. 2.10 The window generated by *ApproximateIntTutor*

Click the "left" button in the Riemann Sums pane, then click the "Display" button to recompute and regraph our function with the new method. Notice that other methods are available in the Riemann Sums pane: "upper," "lower," and "random." These rectangle methods work as follows. Measure the height of the rectangle over the interval $[x_j, x_{j+1}]$ by:

Upper:	choose the maximum of $f(x_j)$ and $f(x_{j+1})$.
Lower:	choose the minimum of $f(x_j)$ and $f(x_{j+1})$.
Random:	randomly choose a point c_j, then use $f(c_j)$.

Experiment with all the methods!
Now let's get a quick comparison of several different methods: click the "Compare..." button. The new window, Fig. 2.11 shows the values given by each of the methods. How do they compare with the actual value given in the original window below the plot?

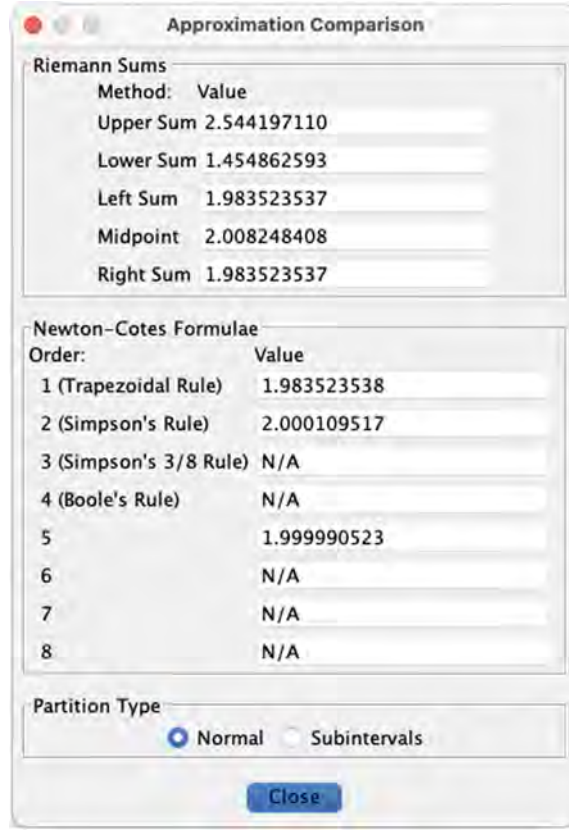


Fig. 2.11 The comparison window generated by *ApproximateIntTutor*

Maple has a very sophisticated integrator. The default method for computing a definite integral is to find an antiderivative, then use the Fundamental Theorem of Calculus to calculate the value. We can tell Maple to directly go to numeric procedures with either

$$evalf(Int(\ldots)) \quad \text{or} \quad int(\ldots, numeric)$$

and bypass trying to find an antiderivative. Maple's default method for numerically integrating is to use the *Numerical Algorithms Group* (NAG) routines[10] which include *Gaussian quadrature*—the next integration method we will study. If the NAG programs results are not satisfactory, then Maple begins with *Clenshaw-Curtis quadrature*[11] that was developed in 1960.

For a fascinating look at how Maple evaluates an integral, first enter

```
> infolevel[int] := 5 :
```

Then enter any integral—start with a really easy one like

```
> int(2x, x = 0..1, numeric);
```

grab a cup of coffee, hit *enter*, and watch Maple work.

For some fun, set *infolevel*[*int*] to 5, then try evaluating the integral of

$$\int_0^1 \frac{e^{-x^2}}{x^2+1}\,dx$$

first without and then with the *numeric* option, but refill that cup of coffee first.

2.3 Gaussian Quadrature

Johann Carl Friedrich Gauss is often called "The Prince of Mathematicians." In 1814, while Gauss was finishing constructing an astronomical observatory, he wrote a paper[12] on approximating integrals. Gauss's technique was studied and extended by Elwin Christoffel in 1858. There are several good ways to develop Gauss's method. We'll use the easiest.

[10] See www.nag.com/numeric/cl/nagdoc_latest/html/d01/d01intro.html.

[11] See Clenshaw and Curtis. "A method for numerical integration on an automatic computer," Numer. Math. 2, 197–205 (1960).

[12] "Methodus nova integralium valores per approximationem inveniendi," *Comment Soc Regiae Sci Gottingensis Recentiores, 1814.*

Table 2.4 Numerical Quadrature Rules as Sums

Left end point	L_n	$= \frac{1}{n} f(x_0) + \frac{1}{n} f(x_1) + \cdots + \frac{1}{n} f(x_{n-1})$
Right end point	R_n	$= \frac{1}{n} f(x_1) + \frac{1}{n} f(x_2) + \cdots + \frac{1}{n} f(x_n)$
Midpoint	M_n	$= \frac{1}{n} f(x_{m_1}) + \frac{1}{n} f(x_{m_2}) + \cdots + \frac{1}{n} f(x_{m_n})$
Trapezoid	T_n	$= \frac{1}{2n} f(x_0) + \frac{1}{n} f(x_1) + \cdots + \frac{1}{n} f(x_{n-1}) + \frac{1}{2n} f(x_n)$
Simpson's rule	S_n	$= \frac{1}{3n} f(x_0) + \frac{4}{3n} f(x_1) + \frac{2}{3n} f(x_2) + \cdots + \frac{4}{3n} f(x_{n-1}) + \frac{1}{3n} f(x_n)$

In order to simplify our study of Gauss' work, we'll focus on $\int_{-1}^{1} f(x)\,dx$ since any interval $[a, b]$ can be mapped to and from $[-1, 1]$ using the transformations

$$T(t) = \frac{1-t}{2} \cdot a + \frac{1+t}{2} \cdot b \quad \text{and} \quad T^{-1}(t) = 2 \cdot \frac{t-a}{b-a} - 1.$$

2.3.1 In Search of Improvements

Collect all the rules we've seen so far into Table 2.4 writing them as expanded sums:

Gauss asked the question, "What features do these sums have in common?"

2.3.2 Patterns

Gauss observed that each expression has the form *a sum of constants times function values at nodes*; that is, a *weighted sum of function values*,

$$A_n = w_1 \cdot f(x_1) + w_2 \cdot f(x_2) + \cdots + w_n \cdot f(x_n),$$

with different sets of weights w_i and different sets of nodes x_i at which to evaluate the function.

Then Gauss posed the questions:

- Is there a "best" set of weights $\{w_i\}$?
- Is there a "best" set of nodes $\{x_i\}$?

The answers will depend on what "best" means.

Since we have $2n$ 'unknowns' w_i and x_i, let's look for a set of weights and a set of nodes that gives the exact value for the integral of a $2n-1$ degree polynomial. (*Remember: $2n$ values uniquely determine a $2n-1$ degree polynomial.*) We'll declare these the best $\{w_i\}$ and $\{x_i\}$ for an n point quadrature since all "reasonable" functions can be approximated by polynomials.[13]

Gauss computed the nodes and weights up to order $n = 7$ (to 16 digits) *by hand!*

Determining the Weights and Nodes for $n = 3$

Set $n = 3$. Then $2n - 1 = 5$. Determine the choice of w_i and of x_i so that

$$\int_{-1}^{1} x^p \, dx = \sum_{k=1}^{3} w_k \cdot x_k^p \tag{2.12}$$

exactly for $p = 0, 1, \ldots, 5$.

The range 0 to 5 for the power p in (2.12) above gives us a system of six nonlinear equations:

$$\left\{ \begin{array}{cccc} w_1 & w_2 & w_3 & 2 \\ w_1 x_1 & w_2 x_2 & w_3 x_3 & 0 \\ w_1 x_1^2 & w_2 x_2^2 & w_3 x_3^2 & \frac{2}{3} \\ w_1 x_1^3 & w_2 x_2^3 & w_3 x_3^3 & 0 \\ w_1 x_1^4 & w_2 x_2^4 & w_3 x_3^4 & \frac{2}{5} \\ w_1 x_1^5 & w_2 x_2^5 & w_3 x_3^5 & 0 \end{array} \right\} \tag{2.13}$$

Use your favourite method of solving a nonlinear system to find

$$\left\{ \begin{array}{ccc} x_1 = -\sqrt{\frac{3}{5}}, & x_2 = 0, & x_3 = \sqrt{\frac{3}{5}} \\ w_1 = \frac{5}{9}, & w_2 = \frac{8}{9}, & w_3 = \frac{5}{9} \end{array} \right\}$$

Our Gaussian quadrature for $n = 3$ is

$$G_3(f) = \frac{5}{9} f\left(-\sqrt{\frac{3}{5}}\right) + \frac{8}{9} f(0) + \frac{5}{9} f\left(\sqrt{\frac{3}{5}}\right).$$

[13] Look up the *Weierstrass Approximation Theorem*.

Solving the system (2.13) for values of the weights and nodes is quite easy with the help of a computer algebra system. First create the system of equations.

```
> N := 3 :
  theSystem :=
    [seq(int(x^p, x = -1..1) = sum(w_k · x_k^p, k = 1..N), p = 0..2N - 1)] :
```

Now, solve the system using the "explicit" option to avoid *RootOf* expressions.

```
> solve(theSystem, explicit = true);
```

The *solve* command will return 6 versions of the solution, but they're just permutations of the variables—they are all really the same solution. *Verify this!* Adding an assumption on the order of the x_is will remove the duplicates, but generate some warnings.

```
> solve(theSystem, explicit = true, useassumptions = true)
      assuming x_1 < x_2 < x_3;
    Warning, solve may not respect assumed property 'real' on 'x[1]'.
    Warning, solve may not respect assumed property 'real' on 'x[2]'.
    Warning, solve may not respect assumed property 'real' on 'x[3]'.
```

$$\left\{ w_1 = \frac{5}{9}, w_2 = \frac{8}{9}, w_3 = \frac{5}{9}, x_1 = -\frac{\sqrt{15}}{5}, x_2 = 0, x_3 = \frac{\sqrt{15}}{5} \right\}$$

Try this for $n = 4$, 5, and 6. *The assumptions on order may need to be dropped.*

With the weights w_i and the nodes/knots x_i in hand, it's time for some testing.

2.3.3 Testing Gauss

First, define the weight and node values. Then we'll generate a random 5th degree polynomial to test, followed by a random 7th degree polynomial. By seeding Maple's random number generator with a value by using the *randomize* command, we can reproduce the polynomials.

```
> w, x := [5/9, 8/9, 5/9], [-sqrt(3/5), 0, sqrt(3/5)]
```

```
> randomize(3019758616) : # Replace the value with your phone number.
```

```
> randpoly(t, degree = 5) :
  p := unapply(%, t);
```

$$p := t \mapsto 13 \cdot t^5 + 11 \cdot t^4 + 93 \cdot t^3 - 27 \cdot t^2 - 18 \cdot t + 95$$

```
> Gauss3 := sum(w_i · p(x_i), i = 1..3);
  ExactValue := int(p(t), t = −1..1);
```

$$Gauss3 := \frac{882}{5}$$

$$ExactValue := \frac{882}{5}$$

The values match!

Now let's try a random 7th degree polynomial.

```
> randpoly(t, degree = 7) :
  q := unapply(%, t);
```

$$q := t \mapsto 77 \cdot t^7 - 98 \cdot t^6 - 66 \cdot t^5 - 68 \cdot t^2 - 21 \cdot t + 10$$

```
> Gauss3 := sum(w_i · q(x_i), i = 1..3);
  ExactValue := int(q(t), t = −1..1);
```

$$Gauss3 := -\frac{3664}{75}$$

$$ExactValue := -\frac{160}{3}$$

The values do not match! *Explain why these results are different and why that's not a surprise.*

2.3.4 An Error Bound for Gaussian Quadrature

The error bound for Gaussian quadrature shows how good the technique is compared to the basic numerical integration techniques we have studied. Deriving the error formula is well above the scope of our current investigation, so we will just state the result in a convenient form.

Theorem 2.1 (Gaussian Quadrature Error Bound) *Let f have $2n$ continuous derivatives on $[-1, 1]$. Then for the error*

$$\varepsilon_n = \left| G_n - \int_{-1}^{1} f(x)dx \right|,$$

we have

$$\varepsilon_n \leq \frac{2^{2n+1}(n!)^4}{(2n+1)\big((2n)!\big)^3} \cdot M_{2n}$$

where M_{2n} is a constant chosen so that $M_{2n} \geq \max \left| f^{(2n)}(x) \right|$ for all x in $[-1, 1]$.

This error estimate can be difficult to use as it requires the $2n$-th derivative of $f(x)$—which can be nearly impossible to generate with only data for f available.

A practical error estimate for Gaussian quadrature can be found by comparing the values for the desired definite integral that are produced by two different sets of nodes/weights. For example,

$$\varepsilon_n \approx |G_n - G_{2n}| .$$

See [7, Sect. 4.9] for a complete discussion of this technique.

2.3.5 Values of Gaussian Weights and Nodes

Prior to the 1960s, the weights and nodes were nearly always taken from published tables such as those in Abramowitz and Stegun's *Handbook* noted below. There are numerous online sources listing Gaussian nodes and weights; for example, the data appears:

1. Beginning at page 916 of the classic *Handbook of Mathematical Functions* by Milton Abramowitz and Irene Stegun published by the National Bureau of Standards [9] (now the National Institute of Standards and Technology),
2. Pavel Holoborodko's Numerical Integration,
3. Mike Kamermans' Gaussian Quadrature Weights and Abscissae, or
4. Casio's Ke!san Online Calculator.

Currently, computer algebra systems can produce the nodes and weights quickly. Maple 2022 found the Gaussian nodes and weights for $N = 41$ to 100 digits in approximately $1/10$ th of a second.

We'll calculate the values directly using Carl Jacobi's formulas from his 1826 paper [10]. Set $P_n(x) = \frac{1}{2^n n!} \cdot \frac{d^n}{dx^n}\left[\left(x^2-1\right)^n\right]$ (the Legendre polynomials). Then

$$\{x_i\}_{i=1}^n = \{zeros\ of\ \mathrm{P_n}\} \quad \text{and} \quad w_i = \frac{2}{\left(1-x_i^2\right)\left[P_n'(x_i)\right]^2} \tag{2.14}$$

With a modest amount of work, we can set up the calculations in Maple.[14] Let's find the values for G_5. First, get P_5 from the *orthopoly* package.

```
> with(orthopoly) :
```

```
> p := unapply(P(5, x), x);
```

$$p := x \mapsto \frac{63}{8} \cdot x^5 - \frac{35}{4} \cdot x^3 + \frac{15}{8} \cdot x$$

[14] A worksheet for computing nodes and weights is available from the Maple Cloud: "Calculate Nodes and Weights for Gaussian Quadrature."

The nodes are the roots of p; for example,

```
> X := sort([solve(p(x) = 0.0, x)]);      # Using '= 0.0' returns decimal answers.
```

$$X := [-0.9061798459, -0.5384693101, 0., 0.5384693101, 0.9061798459]$$

Next, compute the weights from the formula for w_i in (2.14) using the collection of nodes that we just found.

```
> W := [seq(2/((1 - X[i]^2) * p'(X[i])^2), i = 1..5)];
```

$$W := [0.2369268850, 0.4786286708, 0.5688888888, 0.4786286708, 0.2369268850]$$

Now for any function $f(x)$, we have

```
> Gauss5 := sum(W[i] * f(X[i]), i = 1..5);
```

Try it on your favourite function!

2.4 Gauss-Kronrod Quadrature

One difficulty in Gaussian quadrature is that increasing the number of nodes requires recomputing all the values of the nodes, weights, and the function evaluations—discarding the previous calculations. That's a lot of new computational work and a lot of old work thrown away!

2.4.1 Aleksandr Kronrod's Idea

In 1964, Aleksandr Kronrod[15] discovered that he could interlace $n + 1$ new nodes with n original Gaussian nodes creating a $2n + 1$ node rule with order $3n + 1$. A $2n + 1$ node Gaussian quadrature would have order $4n + 1$, but would require significant extra computation for an increase of only n in order over Kronrod's method. The new rule does require new weights for all the nodes, but the original function evaluations can be reused—the function evaluations are the most "expensive" (requiring the most work) part computationally.
Bad news: calculating the values of Kronrod's nodes and weights is well beyond the scope of this text, but we can easily https://www.advanpix.com/2011/11/07/gauss-kronrod-quadrature-nodes-weights/#Tabulated_Gauss-Kronrod_weights_and_abscissae

[15] Kronrod, A. S. (1964.) "Integration with control of accuracy" (*in Russian*), *Dokl. Akad. Nauk SSSR* 154, 283–286.

look them up. *The nodes are the roots of the* https://mathshistory.st-andrews.ac.uk/Biographies/Stieltjes/ Stieltjes *or* Stieltjes-Legendre polynomials. *(See Appendix 2,* pg. 70.)

Gauss-Kronrod quadrature is used in Maple, Mathematica, Matlab, and Sage; it's included in the QUADPACK library, the GNU Scientific Library, the NAG Numerical Libraries, and in the R Statistical Computing Environment. Kronrod's method is the basis of numerical integration in TI calculators. *(Casio calculators use Simpson's rule; HP calculators use an adaptive* Romberg *method.)*

2.4.2 Gauss-Kronrod Quadrature in Practice

$GK_{7,15}$ is a widely used implementation based on a Gaussian quadrature with 7 nodes; Kronrod adds 8 interlaced nodes for a total of 15. Table 2.5 list the weights and nodes of G_7 and $GK_{7,15}$.

Table 2.5 Weights and Nodes of Gauss-Kronrod Quadrature $GK_{7,15}$ on $[-1, 1]$

Gauss-7 nodes		Weights
±0.40584 51513 77397		0.38183 00505 05119
±0.74153 11855 99394		0.27970 53914 89277
±0.94910 79123 42759		0.12948 49661 68870
Kronrod-15 nodes		Weights
±0.20778 49550 07898	*K*	0.20443 29400 75298
±0.40584 51513 77397	*G*	0.19035 05780 64785
±0.58608 72354 67691	*K*	0.16900 47266 39267
±0.74153 11855 99394	*G*	0.14065 32597 15525
±0.86486 44233 59769	*K*	0.10479 00103 22250
±0.94910 79123 42759	*G*	0.06309 20926 29979
±0.99145 53711 20813	*K*	0.02293 53220 10529

As with Gaussian quadrature, deriving an error bound for $GK_{7,15}$ is well is beyond our scope, so we will just state the result. Let

$$GK_{7,15} = \sum_{j=1}^{15} w_j \, f(x_j) \quad \text{and} \quad G_7 = \sum_{k=1}^{7} w_k \, f(x_k).$$

Then

$$\varepsilon_{7,15} \approx \left| G_7 - GK_{7,15} \right|.$$

Kahaner, Moler, and Nash[16] say, in practice, to use the error estimate

$$\varepsilon_{7,15} \approx \left(200 \left| G_7 - GK_{7,15} \right| \right)^{3/2}.$$

Since the nodes and function evaluations of G_7 are reused in $GK_{7,15}$, this estimate doesn't require extra work—that makes it very useful and practical. It is easy to write a simple expression to compute $\left| G_7 - GK_{7,15} \right|$ by replacing the Kronrod weights with the difference between the Gaussian and Kronrod weights.

A *GK* Sample

Use Gauss-Kronrod quadrature to find $\int_{-1}^{1} e^{-x^2} dx$ and give an error estimate.

Using Maple for the computations gives:

$$G_7 = \sum_{k=1}^{7} w_k f(x_k) = 1.49364828886941$$

$$GK_{7,15} = \sum_{k=1}^{15} w_k f(x_k) = 1.49364826562485$$

The error estimate is

$$\varepsilon_{7,15} \approx \left| G_7 - GK_{7,15} \right| = 2.324456 \cdot 10^{-8}.$$

Kahaner, Moler, & Nash's estimate of the error is

$$\varepsilon_{7,15} \approx \left(200 \left| G_7 - GK_{7,15} \right| \right)^{3/2} = 1.002368 \cdot 10^{-8}.$$

[16] Kahaner, Moler, & Nash, *Numerical Methods and Software*, Prentice-Hall, 1989.

Use Maple's *int* command with the *numeric* option.

```
> int(f(x), x = -1..1, numeric);
                    1.49364826562485
```

This answer matches $GK_{7,15}$.

The default first method that Maple attempts for numerical integration is NAG's "d01ajc",[17] an adaptive $GK_{10,21}$ rule. See Maple's Online Help for evalf/int to see the list of methods that are available. TI calculators use an adaptive $GK_{7,15}$ for their `fnInt` function.

2.5 Transforming the Integral Over $[a, b]$ to $[-1, 1]$

The process of changing the interval in a definite integral is not difficult, but does require us to remember the Chain Rule since we are performing a change of variables substitution in the integral.

Transforming the domain $[-1, 1]$ to $[a, b]$ linearly can be done with

$$\begin{aligned} T_{[a,b]}(t) &= \frac{1-t}{2} \cdot a + \frac{1+t}{2} \cdot b \\ &= \frac{b-a}{2} \cdot t + \frac{b+a}{2}, \end{aligned}$$

and then

$$T'_{[a,b]}(t) = \frac{b-a}{2}.$$

Now, use the substitution $x = T_{[a,b]}(t)$ to remap the integral:

$$\begin{aligned} \int_a^b f(x)\,dx &= \int_{-1}^{1} f\left(T_{[a,b]}(t)\right) \cdot T'_{[a,b]}(t)\,dt \\ &= \frac{b-a}{2} \cdot \int_{-1}^{1} f\left(\frac{b-a}{2} \cdot t + \frac{b+a}{2}\right) dt \end{aligned}$$

We can apply Gaussian or Gauss-Kronrod quadrature to the transformed integral to determine the approximate value of the original.

[17] Also described on Maple's Help page for NAG[d01ajc].

A Transformed Integral

Let $f(x) = \tan(\sin(x))$ and compute $\int_0^{\pi} f(x)\,dx$. This function has no elementary antiderivative, so we must use numerical methods. Try Gaussian quadrature with 7 nodes in Maple.

```
> GaussNodes := [−0.949108, −0.741531, −0.405845, 0.0, 0.405845, 0.741531,
    0.949108] :
  GaussWeights := [0.129485, 0.279705, 0.381830, 0.417959, 0.381830,
    0.279705, 0.129485] :
```

```
> G7 := y → sum(eval(y, x = GaussNodes_k) · GaussWeights_k, k = 1..7) :
```

Our transformation function T is:

```
> T := t → 1/2 π t + 1/2 π :
```

So the substitution gives

$$\int_0^{\pi} \tan(\sin(x))\,dx = \frac{\pi}{2} \cdot \int_{-1}^{1} \tan\left(\sin\left(\frac{1}{2}\pi\, t + \frac{1}{2}\pi\right)\right) dt.$$

Finally, Maple gives

```
> Int(f(x), x = 0..Pi) ≈ π/2 · G7 (f(T(t))) ;
```

$$\int_0^{\pi} \tan(\sin(x))\,dx \approx 2.66532$$

Compare our result with Maple's default

```
> int(tan(sin(x)), x = 0..Pi, numeric);
                          2.664281971
```

Not bad for only 7 nodes!

2.6 Easy, but Hard: A Class Quadrature Project

Let's work through the computation of an integral that is geometrically quite easy, but requires care when calculating its value numerically.

"Think and Share" Activity

Consider the periodic function $f(x) = \pi x - \lfloor \pi x \rfloor$. We are going to calculate

$$\mathcal{I}_1 = \int_0^{\pi} f(x)\,dx.$$

1. First, graph the function and find the exact value of the integral $\mathcal{I}$ geometrically.
2. Set $n = 12$. Find

a. The exact value	
b. Left endpoint approximation	
c. Right endpoint approximation	
d. Midpoint approximation	
e. Trapezoid rule approximation	
f. Simpson's rule approximation	
g. Gauss-7 quadrature	
h. Gaussian-Kronrod 7-15 quadrature	

3. Investigate how far the error bound estimates are from the true error.
4. Redo the computations above over $[0, 2\pi]$.
5. Determine a formula for the value of $\mathcal{I}_N = \int_0^{N\pi} f(x)\,dx$ for any $N \in \mathbb{N}$.

2.7 Next Steps

There are several directions for further study that we are ready to investigate. Each item in the following list would make a good project to report on to your class.

- N point open or closed Newton-Cotes formulas
- Gaussian methods
  - Gauss-Chebychev
  - Gauss-Legendre
  - Gauss-Laguerre
  - Gauss-Hermite
- Romberg integration
- Adaptive quadrature
- Clenshaw-Curtis quadrature
- Investigate G.W. Bickley's categorization of numerical integration methods in his paper "Formulæ for Numerical Integration" [11].

Exercises

In Exercises 1 to 5, investigate the integrals using: left endpoint, midpoint, trapezoid, Simpson's rule, and a Gaussian quadrature with 3 nodes. Recall, the Gauss 3 nodes and weights are

$$[x_k] = \left[-\sqrt{\frac{3}{5}}, 0, \sqrt{\frac{3}{5}}\right], \quad [w_k] = \left[\frac{5}{9}, \frac{8}{9}, \frac{5}{9}\right].$$

Estimate the error in each case.

1. The integral of the Chebyshev polynomial of

   a. degree 5
   $$\int_{-1}^{1} 16x^5 - 20x^3 + 5x \, dx,$$

   b. degree 7
   $$\int_{-1}^{1} 64x^7 - 112x^5 + 56x^3 - 7x \, dx.$$

   (First calculate the exact values of the integrals.)

2. $cs = \int_0^{2\pi} \cos(\sin(\theta)) \cdot \cos(\theta) \, d\theta.$
3. $S = \int_0^{\pi} \sin(t^2) \, dt.$
4. Lyness' integral $I(\lambda) = \int_1^2 \frac{0.1}{(x-\lambda)^2 + 0.01} \, dx$ for $\lambda = \pi/2$.
5. The Modified Piessens' integral $\int_{-1}^{1} \left| x^2 - \frac{\pi^2}{16} \right|^{0.1} dx.$
6. Prove:
   A trapezoid sum is the average of a left endpoint and a right endpoint sum.
7. Prove:
   Simpson's rule is the weighted average $2/3$ *midpoint sum plus* $1/3$ *trapezoid sum.*
8. Write Simpson's rule as a weighted average of the three sums: left endpoint, midpoint, and right endpoint.
9. Prove that Simpson's rule must integrate any cubic polynomial exactly over any interval $[a, b]$. Using Maple:

   a. Define
   $$p(x) = p_3x^3 + p_2x^2 + p_1x + p_0$$
   to be an arbitrary cubic polynomial and define its quadratic interpolant

$$q(x) = p(a) \cdot \frac{x-m}{a-m} \cdot \frac{x-b}{a-b} + p(m) \cdot \frac{x-a}{m-a} \cdot \frac{x-b}{m-b} + p(b) \cdot \frac{x-a}{b-a} \cdot \frac{x-m}{b-m}$$

where $m = (a+b)/2$.

b. Define $r(x) = p(x) - q(x)$.
c. Verify $r(a)$, $r(m)$, and $r(b)$ all equal zero. So $r(x) = k(x-a)(x-m)(x-b)$.
d. Integrate $\int_a^m r(x)\,dx$ and $\int_m^b r(x)\,dx$.
e. Explain how you can now conclude that Simpson's rule integrates cubic polynomials exactly.

10. Determine an optimal number of panels n for computing the integral

$$\int_0^1 \frac{1}{1+x}\,dx$$

as we did in *Balancing Roundoff Error and* Δx (page 46) with

a. left endpoint sums
b. trapezoid sums,
c. Simpson's rule.

11. Determine an optimal number of panels n for computing the integral

$$\int_{-2}^2 e^{-x^2}\,dx$$

as we did in *Balancing Roundoff Error and* Δx (page 46) with

a. left endpoint sums
b. trapezoid sums,

12. The arclength of a sine curve from $x = a$ to $x = b$ is given by the *elliptic integral*

$$\mathcal{L}(a,b) = \int_a^b \sqrt{1+\cos^2(t)}\,dt.$$

Elliptic integrals have no elementary antiderivative and can only be evaluated using numerical techniques.

a. Determine $\mathcal{L}(0,\pi)$ using Simpson's rule and $GK_{7,15}$. Compare the errors.
b. Determine $\mathcal{L}(0,2\pi)$ using Simpson's rule and $GK_{7,15}$. Compare the errors.

13. Investigate Gaussian and Gauss-Kronrod quadrature (after transforming the interval to $[-1,1]$) of the integral

$$\int_1^2 \frac{10^{-4}}{\left(x-\frac{\pi}{2}\right)^2+10^{-8}}\,dx.$$

This integral has a sharp peak at $x = \pi/2$. Compare your results to those in Table 2.2 (pg. 51).

14. Use the methods of "Determining the Weights and Nodes for $n = 3$" (pg. 56) to calculate the Gaussian quadrature nodes and weights for

    a. $n = 5$,
    b. $n = 6$,
    c. $n = 7$,
    d. $n = 21$.

15. Explain why the integrals below with Legendre polynomials p_k and Stieltjes polynomial E_k

$$\int_{-1}^{1} p_7(x)\,dx, \qquad \int_{-1}^{1} E_8(x)\,dx, \qquad \int_{-1}^{1} p_7(x)\cdot E_8(x)\,dx$$

    are all zero. Use numerical methods to approximate each.

16. Write a Maple function/program implementing:

    a. Left- and right-endpoint sums
    b. Midpoint and trapezoid sum
    c. Random-point Riemann sum
       One way to generate a random point in $[x_k, x_{k+1}]$ with Maple is:
       define $p := rand(0..1.0)$, then $c_k := x_k + p() \cdot \Delta x$.
    d. Simpson's rule
    e. Gauss G_7
    f. Gauss-Kronrod $GK_{7,15}$

    Test your procedures with several integrals.

17. Define the sums

$$G_7 = \sum_{k=1}^{7} w_{k,G} \cdot f(x_k) \quad \text{with Gauss-7 nodes and Gauss-7 weights,}$$

$$G_K = \sum_{k=1}^{7} w_{k,K} \cdot f(x_k) \quad \text{with Gauss-7 nodes and Kronrod-(7,15) weights,}$$

$$K_{15} = \sum_{k} w_{k,K} \cdot f(x_{k,K-G}) \quad \text{with Kronrod non-Gauss-7 nodes and Kronrod-(7,15) weights.}$$

    Write the quadrature $GK_{7,15}$ and the error $\varepsilon_{7/15}$ using only G_7, G_K, and K_{15}.

18. Prove the inequalities below hold for continuous functions for Maple's "extra" Riemann Sums methods upper, lower, and random.

    a. $\textit{lower sum} \le \int_a^b f(x)\,dx \le \textit{upper sum}$
    b. $\textit{lower sum} \le \textit{random sum} \le \textit{upper sum}$

19. Research and report on *adaptive quadrature*.
20. Research and report on *Romberg integration*.
21. Research and report on *Clenshaw-Curtis quadrature*.
22. Research and report on *two dimensional Gaussian quadrature*.
23. Choose three test integrals from the *Menagerie of Test Integrals* (pg. 69). Compare the calculators' methods on these integrals and estimate the error.

    a. Casio using Simpson's rule,
    b. HP using an adaptive Romberg integration, and
    c. TI using Gauss-Kronrod $GK_{7,15}$.

Appendices

2.8 A Menagerie of Test Integrals

Integrals for Testing Numerical Quadratures[18]

Lyness:

1. $I(\lambda) = \displaystyle\int_1^2 \frac{0.1}{(x-\lambda)^2 + 0.01}\,dx$

Piessens, de Doncker-Kapenga, Überhuber, & Kahaner:

1. $\displaystyle\int_0^1 x^\alpha \log\left(\frac{1}{x}\right)\,dx = \frac{1}{(1+\alpha)^2}$
2. $\displaystyle\int_0^1 \frac{4^{-\alpha}}{\left(x-\frac{\pi}{4}\right)^2 + 16^{-\alpha}}\,dx$
   $= \tan^{-1}\left((4-\pi)4^{\alpha-1}\right)$
   $+ \tan^{-1}\left(\pi 4^{\alpha-1}\right)$
3. $\displaystyle\int_0^\pi \cos(2^\alpha \sin(x))\,dx = \pi\, J_0(2^\alpha)$
4. $\displaystyle\int_0^1 \left|x-\tfrac{1}{3}\right|^\alpha\,dx = \frac{\left(\frac{2}{3}\right)^{\alpha+1} + \left(\frac{1}{3}\right)^{\alpha+1}}{1+\alpha}$
5. $\displaystyle\int_0^1 \left|x-\frac{\pi}{4}\right|^\alpha\,dx$
   $= \frac{(1-\frac{\pi}{4})^{\alpha+1} + (\frac{\pi}{4})^{\alpha+1}}{1+\alpha}$
6. $\displaystyle\int_{-1}^{+1} \frac{1}{\sqrt{1-x^2}}\,\frac{1}{1+x+2^{-\alpha}}\,dx$
   $= \frac{\pi}{\sqrt{(1+2^{-\alpha})^2-1}}$

[18] This set of integrals appears in D. Zwillinger's *Handbook of Integration* p. 272, (A K Peters/CRC Press, 1992).

7. $\int_0^{\pi/2} \sin^{\alpha-1}(x)\,dx = 2^{\alpha-2}\,\frac{\Gamma^2(\frac{\alpha}{2})}{\Gamma(\alpha)}$

8. $\int_0^1 \log^{\alpha-1}\left(\frac{1}{x}\right)\,dx = \Gamma(\alpha)$

9. $\int_0^1 \frac{\cos(2^\alpha x)}{\sqrt{x(1-x)}}dx = \pi\cos(2^{\alpha-1})\,\mathrm{J}_0(2^{\alpha-1})$

10. $\int_0^\infty x^2 e^{-2^{-\alpha}x}\,dx = 2^{3\alpha+1}$

11. $\int_0^\infty \frac{x^{\alpha-1}}{(1+10x)^2}\,dx = \frac{(1-\alpha)\pi}{10^\alpha\sin(\pi\alpha)}$

Berntsen, Espelid, & Sørevik:

13. $\int_0^1 \left|x-\frac{1}{3}\right|^{-1/2}\,dx$ *(singularity)*

14. $\int_{-1}^1 U(x)\,e^{x/2}\,dx$ *(discontinuity)*

15. $\int_0^1 e^{2|x-1/3|}\,dx$ $(C_0\,function)$

16. $\int_1^2 \frac{10^{-4}}{\left(x-\frac{\pi}{2}\right)^2+10^{-8}}\,dx$ *(sharp peak)*

2.9 Legendre and Stieltjes Polynomials for $GK_{7,15}$

The Polynomials of $GK_{7,15}$

The Gaussian nodes for G_7 are the roots of the Legendre polynomial p_7,

$$p_7(x) = \frac{429}{16}x^7 - \frac{693}{16}x^5 + \frac{315}{16}x^3 - \frac{35}{16}x.$$

The additional nodes Kronrod adds for $GK_{7,15}$ are the roots of the Stieltjes polynomial E_8. This polynomial is found by solving the "orthogonality condition" system of equations

$$\mathcal{OC} = \left\{\int_{-1}^1 p_7(x)E_8(x)\,x^k\,dx = 0; \quad \text{for} k = 0..7\right\}$$

for E_8 which gives

$$E_8(x) = c\left[\tfrac{4854324041}{52932681}x^8 - \tfrac{1142193892}{5881409}x^6 + \tfrac{765588166}{5881409}x^4 - \tfrac{501576364}{17644227}x^2 + 1\right].$$

Figure 2.12 shows the roots interlacing.

The Kronrod weights can be found in various ways. See Dirk Laurie's paper "Calculation of Gauss-Kronrod quadrature rules" [12] for an efficient method using Linear Algebra techniques to calculate the weights and nodes. Also read Sotirios Notaris's paper "Gauss-Kronrod Quadrature Formulæ–A Survey of Fifty Years of Research" [13] for the history and for more information on these computations.

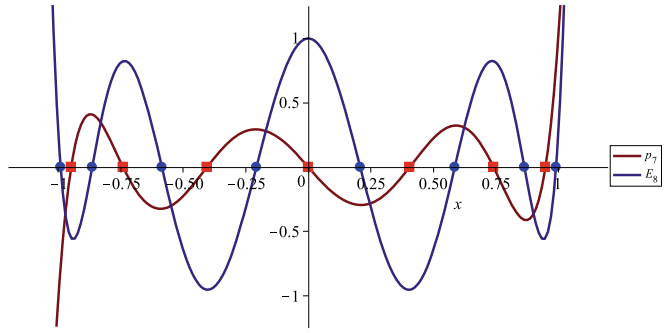


Fig. 2.12 The Polynomials p_7 and E_8

2.10 Links and Others

There are large numbers of websites with information about numerical quadrature. A Google search returns over 6.5 million pages in less than half a second. A Google Scholar search returns over half a million pages in less than a tenth of a second. A Bing search returns over 53 million pages in less than half a second.

- Web searches for "numerical quadrature":
  - Google
  - Google Scholar
  - Bing
- Online integration methods:
  - www.hvks.com/Numerical/webintegration.html

  More information:
  Wikipedia entries:
  - Newton-Cotes formulas
  - Romberg's method
  - Clenshaw-Curtis integration
  - Cubature

  See also: MathWorld (Wolfram Research) and Interactive Educational Modules in Scientific Computing (U of I)
  Investigate:
  - Boole's Rule
  - adaptive quadrature
  - orthogonal polynomials
  - Vandermonde Matrix
  - Maple's `evalf/Int` command

- The Maple command `ApproximateInt` in the `Student[Calculus1]` package
- Matlab's `integral` command
- Cubature.

References

1. Davis, P.J.: Interpolation and Approximation. Dover Pub., Inc. (1975)
2. Goldberg, D.: What every computer scientist should know about floating-point arithmetic. ACM Comput. Surv. **23**(1), 5–48 (1991). URL https://docs.oracle.com/cd/E19957-01/800-7895/800-7895.pdf
3. Bauldry, W.C.: An Introduction to Computational Mathematics: An Outline. CRC Press (2023)
4. Goldstine, H.H.: A History of Numerical Analysis from the 16th through the 19th Century, vol. 2. Springer Science & Business Media (2012)
5. Epperson, J.F.: An Introduction to Numerical Methods and Analysis. John Wiley & Sons (2021)
6. Talman, L.A.: Simpson's rule is exact for quintics. The American Mathematical Monthly **113**(2), 144–155 (2006)
7. Davis, P., Rabinowitz, P., Rheinbolt, W.: Methods of Numerical Integration. Computer Science and Applied Mathematics. Elsevier Science (2014). URL https://books.google.com/books?id=mbLiBQAAQBAJ
8. Berntsen, J., Espelid, T.O., Sørevik, T.: On the subdivision strategy in adaptive quadrature algorithms. Journal of Computational and Applied Mathematics **35**(1-3), 119–132 (1991)
9. Abramowitz, M., Stegun, I.A.: Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Applied Mathematics Series. National Bureau of Standards (1965)
10. Jacobi, C.G.J.: Üeber gauss neue methode, die werthe der integrale näherungsweise zu finden. Jour. Reine Angew. Math. (1), 301–307 (1826)
11. Bickley, W.G.: Formulae for numerical integration. The Mathematical Gazette **23**(256), 352–359 (1939)
12. Laurie, D.: Calculation of gauss-kronrod quadrature rules. Mathematics of Computation **66**(219), 1133–1145 (1997)
13. Notaris, S.E.: Gauss-kronrod quadrature formulae–a survey of fifty years of research. Electron. Trans. Numer. Anal **45**, 371–404 (2016)

Projects

3

The following presents a collection of student projects involving the topics we've investigated. These projects can be assigned as extended individual work or to small groups. The first six look at uses of numerical derivatives. The seventh project forms a bridge from derivative to integral using Euler's method. The next six projects are based on numerical integration. The last project is a theoretical look at the integral power rule and the indefinite integral $\int x^{-1}\,dx$.

3.1 Polynomials and Roots from Data Using Numerical Derivatives

3.1.1 Monic Polynomials and Their Derivatives

Let's begin with a third degree, monic polynomial

$$q(x) = x^3 + a_2x^2 + a_1x + a_0.$$

Remember that $a_k = q^{(k)}(0)/k!$.

Suppose we know three values of q: the values $q(0)$ and $q(\pm x_1)$ from

$$\mathcal{D} = \{q(0) = y_0,\ x_1,\ q(x_1) = y_1^+,\ q(-x_1) = y_1^-\}.$$

Then since $q(0) = a_0$, we immediately have $a_0 = y_0$.

Let's use a difference formula to determine a_1.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

W. C. Bauldry, *Computational Calculus*, Synthesis Lectures on Mathematics & Statistics,
https://doi.org/10.1007/978-3-031-29658-1_3

1. Calculate
$$\frac{q(x_1) - q(-x_1)}{x_1 - (-x_1)} = \frac{y_1^+ - y_1^-}{2x_1}.$$
Use the formula for q on the left side to compute the value in terms of x_1 formally; use the given data on the right side.
2. Adjust the result to give a value for just a_1.

Use a second difference formula to determine a_2.

1. Calculate
$$\frac{q(x_1) - 2q(0) + q(-x_1)}{2x_1^2} = \frac{y_1^+ - 2y_0 + y_1^-}{2x_1^2}.$$
Again use the formula for q on the left side to compute the value formally; use the given data on the right side.
2. Adjust the result to give a value for just a_2.

3.1.2 The Roots

3.1.2.1 Depressed Cubics

Translate by $x = t - a_2/3$. Notice what the translation did to the x^2 term. The depressed polynomial's roots are all translated by $a_2/3$ from the original roots, so it's quite easy to recover the original function's roots.

This technique will depress any polynomial's second-highest order term. For $p(x) = a_n x^n + a_{n-1}x^{n-1} = \ldots$, translate with $x = t - a_{n-1}/(n\,a_n)$ to eliminate the x^{n-1} term.

3.1.2.2 Roots in Terms of Coefficients

Use the formulas developed above to replace the coefficients with the values from those in $\mathcal{D}$.

3.1.3 The Project

Use your formulas from above to determine the monic cubic polynomial and its roots for randomly generated data for $\mathcal{D}$ from Maple.

```
> randomize(8282623050) :      # replace with your phone number
```

```
> xr, yr := rand(0.0..9.0), rand(−5.0..5.0) :
```

$$> \mathcal{D} := [yr(), xr(), yr(), yr()]; \qquad \text{\# where } \mathcal{D} := [y_0, x_1, y_1^+, y_1^-],$$

Graph your polynomial and show its roots.

See [1] for an extension of these ideas to quartics.

3.2 Space Shuttle Acceleration

3.2.1 Space Shuttle Acceleration: The Situation

The Space Shuttle[1] had three phases from ignition to achieving orbit. The first phase began by the shuttle lifting off the launch pad using solid rocket boosters (SRBs) to accelerate extremely quickly. At approximately two minutes, the SRBs, having burned approximately 500,000 kg of solid fuel, separated from the shuttle falling back to Earth.

The second phase began with SRB separation and lasted approximately 6.5 m. The shuttle speed increased to 17,500 mph—the speed needed to achieve orbit. (This speed is good deal less than Earth's escape velocity of 25,000 mph.)

The third phase began at about 9 m when the external fuel tank was jettisoned and the shuttle entered orbit.

Table 3.1 lists NASA's data for STS-121 from July 4, 2006.

Table 3.1 Space shuttle mission STS-121 data

Time	Altitude	Velocity	Time	Altitude	Velocity	Time	Altitude	Velocity
0	0	0	180	77,485	1634	360	108,543	3860
20	1,244	139	200	85,662	1800	380	107,690	4216
40	5,377	298	220	92,481	1986	400	106,539	4630
60	11,617	433	240	98,004	2191	420	105,142	5092
80	19,872	685	260	102,301	2417	440	103,775	5612
100	31,412	1026	280	105,321	2651	460	102,807	6184
120	44,726	1279	300	107,449	2915	480	102,552	6760
140	57,396	1373	320	108,619	3203	500	103,297	7327
160	67,893	1490	340	108,942	3516	520	105,069	7581

[1] Adapted from NASA's "*Space Shuttle Ascent*" student project.

3.2.2 The Project

1. Using different divided difference formulas, compute and graph the shuttle's acceleration. Compare the results from each method.
2. Determine the maximum acceleration and its time.
3. Approximate the acceleration when the shuttle entered orbit at 9 m.
4. Explain why the shuttle's velocity cannot be computed from $\Delta Altitude/\Delta time$.

3.3 Measuring Capacitance and Inductance with Derivatives

Capacitance is the ability of a circuit element to collect and store an electrical charge. Capacitors are used to smooth oscillating voltages, set the frequency of an oscillator, couple signals together, among many other applications. Inductance is ability of a circuit element to induce a voltage by a variation of current either in the circuit element itself or in a neighboring component. Inductors are used for attenuating or filtering high frequency noise, storing energy, and transferring energy to another circuit, among many other applications.

3.3.1 Capacitance and Inductance Equations

Let C be capacitance in farads, L be inductance in henries, i be current in amperes, and v be voltage in volts.[2] The equation modeling capacitance is

$$i = C \cdot \frac{dv}{dt}. \tag{3.1}$$

The model shows that capacitors block constant voltages (DC) and pass alternating voltages (AC). *Explain how the equation indicates these properties.*

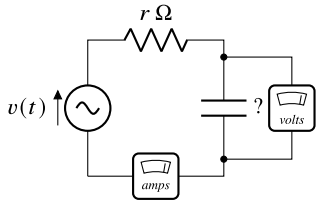
The equation modeling inductance is

$$v = L \cdot \frac{di}{dt}. \tag{3.2}$$

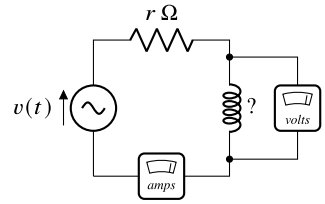
The model shows that inductors resist constant currents and pass alternating currents. *Explain how the equation indicates these properties.*

Notice the complementary forms of the modeling equations.

[2] Both farads and henries are much too large for practical circuits. Real circuits most often use microfarads (10^{-6} farads) or picofarads (10^{-12} farads) and millihenries (10^{-3} henries). Similar scaling is used for voltage and current. Always pay close attention to the units.



(a) The Capacitor Tester



(b) The Inductor Tester

Fig. 3.1 Simple testing circuits

Combining capacitors and inductors creates *tuned circuits* that form the basis of radio transmitters and receivers.

3.3.2 The Project: Measuring Component Values

A laboratory has collected the data listed in Table 3.2. *Pay attention to the units!*

1. Using Table 3.2a and (3.1), determine the value of the capacitor being tested.
2. Using Table 3.2b and (3.2), determine the value of the capacitor being tested.
3. Using error estimates for your calculations, give a *tolerance rating* in the form $\pm T\%$ for the tested components.

3.4 Marginal Revenue From Data, Cost, and Profit

3.4.1 Marginal Revenue and Marginal Cost

Marginal revenue is defined to be the change in revenue r from selling an additional item, or the derivative dr/dq when the change in quantity q sold becomes arbitrarily small. Marginal cost is defined to be the change in cost c from producing an additional item, or again, the derivative when the change in quantity produced becomes arbitrarily small. With revenue equal to price times quantity sold, profit is given by

$$Profit = Price \times Quantity - Cost.$$

A little basic calculus shows that profit is maximized when marginal revenue equals marginal cost; i.e., $dr/dq = dc/dq$. *Verify this!*

Table 3.2 Capacitor and Inductor Data

(a)The capacitor data			(b) The inductor data		
Time	Volts	Milliamps	Time	Volts	Milliamps
0.01	104.60	−6.41	0.01	89.00	4810.0
0.02	−64.67	−16.78	0.02	104.60	15280.0
0.03	−64.56	16.79	0.03	33.94	22770.0
0.04	104.70	6.36	0.04	−64.76	21110.0
0.05	−0.13	−20.74	0.05	−110.00	11670.0
0.06	−104.60	6.41	0.06	−64.65	2228.0
0.07	64.24	16.84	0.07	33.96	570.0
0.08	64.62	−16.78	0.08	104.70	8066.0
0.09	−104.80	−6.30	0.09	88.88	18530.0
0.10	0.05	20.74	0.10	−0.13	23340.0
0.11	104.40	−6.52	0.11	−89.17	18500.0
0.12	−64.70	−16.77	0.12	−104.60	8060.0
0.13	−64.16	16.84	0.13	−33.18	580.0
0.14	104.60	6.40	0.14	65.13	2265.0
0.15	−0.62	−20.74	0.15	110.00	11700.0
0.16	−104.60	6.42	0.16	64.62	21110.0
0.17	65.16	16.70	0.17	−33.77	22780.0
0.18	63.69	−16.91	0.18	−104.80	15330.0
0.19	−104.80	−6.29	0.19	−88.80	4784.0
0.20	0.10	20.74	0.20	0.049	0.0

3.4.2 The Project

1. Choose a grain to model: soybeans, oats, corn, or barley.
2. Use the data presented in Tables 3.3 and 3.4 to generate revenue and cost functions.
3. Numerically differentiate your functions to find marginal revenue and marginal cost.
4. Determine optimal profit points.

Table 3.3 Historical commodity selling prices

Year	Price Per Bushel ($)				Year	Price Per Bushel ($)			
	Soybeans	Oats	Corn	Barley		Soybeans	Oats	Corn	Barley
2001	4.13	1.67	1.75	2.00	2011	11.40	3.55	6.05	5.45
2002	5.33	1.99	2.17	2.51	2012	14.10	3.80	7.05	5.70
2003	6.96	1.43	2.28	2.37	2013	12.40	3.60	4.10	5.00
2004	5.58	1.49	1.82	2.02	2014	9.40	2.95	3.25	3.50
2005	5.39	1.58	1.79	2.11	2015	8.40	2.30	3.25	3.75
2006	6.03	2.08	2.88	2.51	2016	9.10	1.95	3.20	–
2007	9.60	2.87	4.17	4.55	2017	8.90	2.65	3.00	–
2008	9.65	2.68	3.78	5.06	2018	8.05	2.45	3.30	3.65
2009	9.05	2.15	3.40	3.05	2019	8.40	2.90	3.70	3.85
2010	11.20	2.90	5.10	3.00	2020	10.60	2.55	4.05	4.30

Source "Historical Commodity Prices," USDA

Table 3.4 Cost of production

		Soybeans	Oats	Corn	Barley
A	Operating Costs				
	Seed & Treatment	$97.10	$35.00	$99.20	$28.00
	Fertilizer	$52.92	$133.49	$256.08	$149.91
	Pesticide	$19.33	$23.25	$41.13	$53.75
	Fuel	$25.20	$35.45	$35.97	$31.07
	Drying & Other Costs	$17.75	$17.75	$52.75	$17.75
	Storage Costs	$5.87	$18.77	$22.03	$12.40
B	Fixed Costs				
	Land Taxes	$17.50	$17.50	$17.50	$17.50
	Machinery Costs	$68.01	$68.01	$76.17	$68.01
C	Target Yield				
	Bushels per acre	51.2	110.7	171.4	80.6

Sources "Estimates of Field Crop Production—2020," Markets & Statistics, Manitoba Commission on Agriculture; "Guidelines For Estimating Crop Production Costs—2022," Manitoba Commission on Agriculture

3.5 Estimating Speed and Acceleration from GPS Data

The *Global Positioning System*[3] (GPS) provides latitude, longitude, and altitude data with a time-stamp at a receiver's location by comparing signals from at least four different GPS satellites. We will use a small part of driver's GPS data set to estimate speed (derivative of position) and acceleration (second derivative of position).

3.5.1 The Global Positioning System and Distance

GPS provides latitude and longitude in degrees and time as UTC (Coordinated Universal Time). Our first task is convert subsequent pairs of [*latitude*, *longitude*] to distance traveled. Taking the earth as a sphere of radius $r = 6{,}378 \cdot 10^3$ m (mean radius) will not introduce too much error as we are dealing with small distance increments. Sinnot's Haversine formula [2] offers a convenient method of calculating the distance between two latitude-longitude pairs. The version of the Haversine formula we use ignores elevation changes, but again, our increments are very small, so this does not introduce too much error. The Haversine formula for two latitude-longitude pairs (lat_1, lon_1), (lat_2, lon_2) measured in radians is

$$\begin{aligned}(\Delta_{lat}, \Delta_{lon}) &= (lat_2, lon_2) - (lat_1, lon_1)\\ a &= \sin^2(0.5 \cdot \Delta_{lat}) + \cos(lat_1) \cdot \cos(lat_2) \cdot \sin^2(0.5 \cdot \Delta_{lon})\\ distance &= 2R \cdot \arcsin\big(\min(1, \sqrt{a}\,)\big) = 2R \cdot \arctan(\sqrt{a}, \sqrt{1-a})\end{aligned}$$

where R, the radius of the earth in meters, is

$$R = \begin{cases} 6378 \cdot 10^3 & \textit{uncorrected for latitude}\\ \big(6378 - 21 \cdot \sin(\max(lat_1, lat_2))\big) \cdot 10^3 & \textit{corrected}\end{cases}.$$

3.5.2 The Project

A driver's data appears in Table 3.5.

1. Convert the latitude-longitude pairs in Table 3.5 from degrees to radians.
2. Find the distance between every fifth point.
3. Using different divided difference formulas, compute and graph the driver's speed and acceleration.
4. Compare the results from each method.
5. Estimate the effect changing altitude has on your results.

[3] See www.gps.gov for more information about GPS and for technical documentation.

Table 3.5 The driver's GPS data

n	Lat (deg)	Lon (deg)	Alt (m)	Time (h:m:s)
1	36.234178	–81.659315	988.663742	13:42:00
2	36.234288	–81.659220	987.081672	13:42:01
3	36.234385	–81.659093	986.422759	13:42:02
4	36.234466	–81.658941	985.220085	13:42:03
5	36.234529	–81.658780	982.499996	13:42:04
6	36.234587	–81.658603	981.883204	13:42:05
7	36.234649	–81.658421	981.241166	13:42:06
8	36.234729	–81.658236	979.445309	13:42:07
9	36.234814	–81.658054	979.000861	13:42:08
10	36.234915	–81.657879	978.099293	13:42:09
11	36.235042	–81.657715	976.419456	13:42:10
12	36.235184	–81.657567	975.721942	13:42:11
13	36.235333	–81.657446	974.964100	13:42:12
14	36.235483	–81.657352	972.460857	13:42:13
15	36.235631	–81.657274	970.565406	13:42:14
16	36.235784	–81.657202	969.693777	13:42:15
17	36.235931	–81.657135	969.144298	13:42:16
18	36.236071	–81.657073	968.415241	13:42:17
19	36.236214	–81.657013	968.565132	13:42:18
20	36.236347	–81.656953	968.617595	13:42:19
21	36.236482	–81.656895	968.788547	13:42:20
22	36.236614	–81.656833	968.768365	13:42:21
23	36.236747	–81.656773	968.758979	13:42:22
24	36.236884	–81.656713	968.598957	13:42:23
25	36.237016	–81.656645	968.399636	13:42:24
26	36.237156	–81.656581	967.733557	13:42:25
27	36.237303	–81.656525	966.336469	13:42:26
28	36.237447	–81.656457	965.673376	13:42:27
29	36.237592	–81.656395	965.178910	13:42:28
30	36.237736	–81.656328	964.228664	13:42:29
31	36.237883	–81.656267	963.912202	13:42:30
32	36.238037	–81.656210	963.445788	13:42:31
33	36.238191	–81.656158	962.388462	13:42:32
34	36.238345	–81.656115	962.060205	13:42:33
35	36.238508	–81.656077	961.794292	13:42:34
36	36.238669	–81.656047	961.470454	13:42:35
37	36.238832	–81.656025	973.003082	13:42:36
38	36.238995	–81.656013	971.692902	13:42:37
39	36.239154	–81.656008	971.488434	13:42:38
40	36.239318	–81.656006	970.844879	13:42:39

6. Identify the town where the driver collected the data. Could altitude be a significant factor?

3.6 Fornberg's Algorithm

In 1988, Bengt Fornberg published the paper "Generation of Finite Difference Formulas on Arbitrarily Spaced Grids" [3] developing a method of computing the coefficients of centered difference formulas for a desired derivative to a chosen accuracy. The algorithm is much simpler than most others.

3.6.1 The Algorithm

The algorithm can be programmed without much difficulty. However, pay close attention to each index. The algorithm's *index origin* is 0, but Maple's normal origin is 1. We've modified the algorithm to use equally spaced points as in the formulas we've developed: $x_0, x_0 \pm h, x_0 \pm 2h, x_0 \pm 3h, \ldots$, etc.

Fornberg's Algorithm

Determine the coefficients $\delta_{n,\nu}^m$ for $m = 0, 1, \ldots, M$ of order $\mathcal{O}(h^{n-m+1})$ so that

$$f^{(m)}(x_0) \approx \sum_{\nu=0}^{n} \delta_{n,\nu}^m f(\alpha_\nu).$$

```
Enter M, N
Set α = [−N/2, −N/2 + 1, −N/2 + 2, ..., +N/2]
Set δ^1_{0,0} := 1
Set c1 := 1
for n from 1 to N do
    c2 := 1
    for ν from 0 to n − 1 do
        c3 := a_n − a_ν
        c2 := c2 · c3
        if n < M then δ^n_{n,ν} := 0 end if
        for ν from 0 to min(n, M) do
            δ^m_{n,ν} := (α_n · δ_{n−1,ν} − m · δ^{m−1}_{n−1,ν}) · c3
            next m
        next ν
    for m from 0 to min(n, M) do
        δ^m_{n,n} := (c1/c2) · (m · δ^{m−1}_{n−1,n−1} − α_{n−1} · δ^m_{n−1,n−1})
```

```
        next m
    c1 := c2
    next n
return [δ^j_{M,N} : j = 0..N]
```

3.6.2 The Project

1. Implement Fornberg's algorithm in Maple.
2. Test the algorithm by producing the centered difference formulas that we have already seen.
3. Produce a centered difference formula for the fourth derivative of order $\mathcal{O}(h^6)$.
4. Use your formula above to find $f^{(4)}(x)$ when

   a. $f(x) = \tan(\sin(x))$

   b. $Si(x) = \int_0^x \frac{\sin(t)}{t}\, dt$

   Use plots to compare to the results of analytic differentiation.

3.7 Euler's Method of Solving Differential Equations

Euler's method is the first numerical method learned for solving first-order differential equations (equations with only the first derivative). The method is based on the observation that the tangent line to a function approximates the curve close to the point of tangency.

3.7.1 Euler's Method Graphically

The method is based on the diagram in Fig. 3.2.

Following the tangent line from P_0 takes us to P_1. The smaller the step Δx along the tangent line, the smaller the error for "decently behaved" functions. Euler's method is simply restart and follow the tangent line through P_1 producing P_2. Start again following the tangent line through P_2 producing P_3, and so forth, until reaching the desired ending x-value.

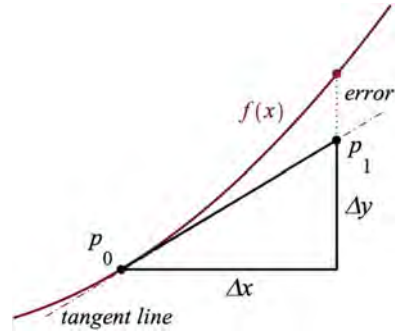


Fig. 3.2 The differential triangle

3.7.2 Euler's Method Algebraically

Euler's method can be derived algebraically starting with the simple formulas

$$\Delta y/\Delta x \approx y' \quad \text{and} \quad y_{\text{next}} = y_{\text{current}} + \Delta y.$$

Combining these relations, we see that

$$y_{\text{next}} \approx y_{\text{current}} + y' \cdot \Delta x.$$

Simply evaluating y' at y_{current} and adding $y' \cdot \Delta x$ to y_{current} brings us to y_{next}. Repeat this procedure to get to the desired point creating the list

$$[x_0, y_0] \to [x_1, y_1^*] \to [x_2, y_2^*] \to \cdots \to [x_n, y_n^*].$$

Figure 3.3 shows the results of Euler's method linking the triangles applied to the differential equation $y' = x - x^4$ with $\Delta x = 0.1$.

3.7.3 The Project

1. Use Euler's method with the with step sizes given below to find $y(2\pi)$ when $y' = \tan(\sin(x))$ with $y(0) = -1.25$.

   a. $h = 2.0$
   b. $h = 1.0$
   c. $h = 0.5$
   d. $h = 0.25$

   Graph each of your solutions and y' together.

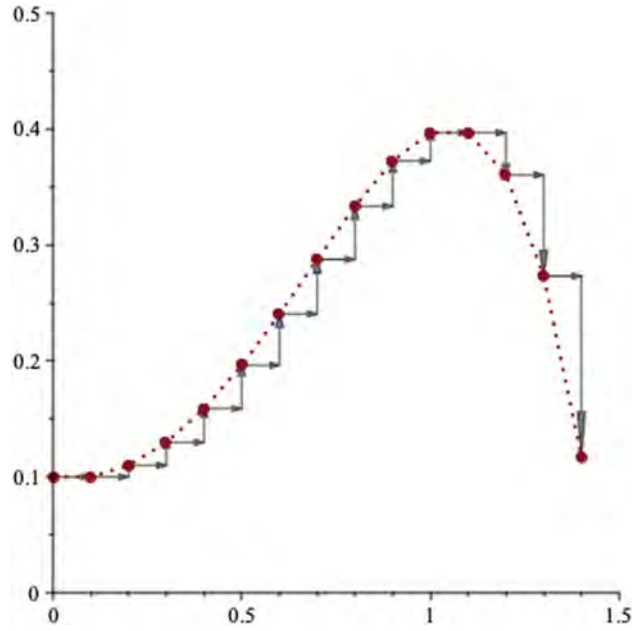


Fig. 3.3 Euler's method in action.eps

2. Find solution to $y' = y \cdot (1 - y)$ with $y(0) = 0.01$ using Euler's method with step size

   a. $h = 2.0$,
   b. $h = 1.0$,
   c. $h = 0.5$,
   d. $h = 0.25$,

   for $0 \leq x \leq 10$. Graph each of your solutions and y' together.

   To check your work graphically with Maple, enter the following:

```
> de1 := y'(x) = tan(sin(x)) :
  de1opts := (y(x), x = 0..2π, y = −2.25..2.25, [y(0) = −1.25], stepsize =
      2.0/n, linecolor = black, method = classical[foreuler]) :

> Explore(DEtools:-DEplot(de1, de1opts), n = 1..6);
```

Then execute:

```
> de2 := y'(x) = y(x) · (1 − y(x)) :
  de2opts := (y(x), x = 0..10, y = −0.1..1.1, [y(0) = 0.01], stepsize =
       2.0/n, linecolor = black, method = classical[foreuler]) :
```

```
> Explore(DEtools:-DEplot(de1, de1opts), n = 1..6);
```

3.8 A Dog's Life

Dogs mature much faster than humans. An old rule of thumb was that one dog year was equivalent to seven human years. However, this rule is a large oversimplification. Just like most mammals, dogs mature much more rapidly when young, then the rate of aging slows as the dog gets older.

3.8.1 Rate of Aging Data

The laboratory at a pet food company has collected data on the aging rate of medium sized dogs for comparing feed rations. The company has released the average over all of their pet foods of their data shown in Table 3.6.

Table 3.6 Average rate of growth for dogs over time

Year	Rate of aging	Year	Rate of aging
0.5	32.17	7.5	2.18
1.0	16.14	8.0	2.13
1.5	10.72	8.5	1.95
2.0	8.07	9.0	1.78
2.5	6.49	9.5	1.85
3.0	5.52	10.0	1.68
3.5	4.68	10.5	1.69
4.0	4.19	11.0	1.55
4.5	3.73	11.5	1.48
5.0	3.34	12.0	1.38
5.5	3.03	12.5	1.37
6.0	2.79	13.0	1.26
6.5	2.47	13.5	1.32
7.0	2.39	14.0	1.25

3.8.2 The Project

The integral

$$Human\ years(T) = \int_0^T rate\ of\ aging(t)\,dt$$

will give a dog's age in human years as a function of time T.

1. Create a graph of *Human years* versus *Dog years* for $T \in [0.5, 14.0]$ by numerically integrating the data of Table 3.6.
2. What type of function will produce the curve you found above?

3.9 Computing the Cost at a Gas Pump

The first gas pumps were invented by Sylvanus Bowser in 1885 and were originally called "Bowsers." The early pumps pumped kerosene into a marked glass cylinder, then drained by gravity to a customer's vessel. Modern gas pumps at service stations are limited to pumping 10 gallons per minute for safety reasons; the rate is set by the U.S. Environmental Protection Agency.[4] Modern fuel pumps are computer controlled. The onboard computer manages pumping, the displays showing data to the customer, and communicates with the dealer's sales system.

3.9.1 Calculating the Cost of the Gas Pumped

How does the pump determine a cost for the customer's purchase at the pump? The pump's computer measures the rate of flow of gas being pumped. Most modern service station pumps use a 4 stroke piston meter that works in a similar fashion to a piston motor. The mechanical motion of the piston is converted to an electrical impulse captured by the computer and measures the rate of flow of gas. The rate of flow, that is, the derivative of the flow function, is integrated to find the volume of gas delivered. Multiplying by price per gallon generates the customer's cost. Thus

$$Flow(t) \quad \Rightarrow \quad Volume = \int_0^T Flow(t)\,dt \quad \Rightarrow \quad Cost = Price \cdot Volume$$

[4] See 61 FR 33033 - Regulation of Fuels and Fuel Additives.

3.9.2 The Project

Calculate the cost of buying 20 gallons of gas from a service station.

1. Visit the web site gasprices.aaa.com/state-gas-price-averages to find the current average gas price P in your county.
2. Use the data given in Table 3.7 to determine the volume of gas pumped by integrating with

Table 3.7 Gas flow measurements

Time (s)	Flow rate (gal/min)
0	8.80
5	7.71
10	9.31
15	8.23
20	8.14
25	6.79
30	8.03
35	7.05
40	6.89
45	7.65
50	6.77
55	7.62
60	7.02
65	8.51
70	7.20
75	8.85
80	8.56
85	9.43
90	8.70
95	8.46
100	7.97
105	9.18
110	9.36
115	6.50
120	7.00

a. an endpoint sum
b. midpoint sum
c. simpson's rule
d. Give an error estimate in the calculated cost.
e. How important is accuracy in this estimation? Calculation speed?

3. Determine the cost of the gas that was pumped.

3.10 The Fourier Power Spectrum

The *Power Spectrum* of a signal f gives the amount of power of a specific frequency component in the signal. The power value at the frequency ω is given by

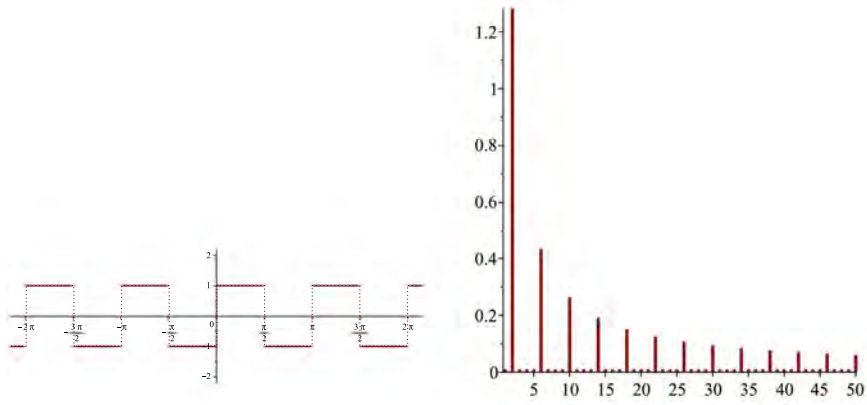
$$P(\omega) = \sqrt{a_\omega^2(f) + b_\omega^2(f)}$$

where $a_\omega(f)$ and $b_\omega(f)$ are the *Fourier coefficients* of the signal f at the frequency ω. The *Spectrum Plot* is a discrete graph that shows the power level at each frequency. Figure 3.4 shows a square wave with its spectrum plot.

The Fourier coefficients a and b are computed with the integrals ($\omega \geq 1$)

$$a_\omega(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(\omega t)\,dt \quad \text{and} \quad b_\omega(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(\omega t)\,dt,$$

and



(a) Square Wave $\mathrm{SW}(t)$ (b) Spectrum Plot of $\mathrm{SW}(t)$

Fig. 3.4 Graph and spectrum graph of $\mathrm{SW}(t) = |\sin(2t)|/\sin(2t)$

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\, dt.$$

Then, when f is differentiable,

$$f(x) = a_0 + \sum_{\omega=1}^{\infty} a_\omega \cos(\omega t) + b_\omega \sin(\omega t).$$

Fourier thought he could expand any function this way. He—rather immodestly—declared that "The theory…will hereafter form one of the most important branches of general physics."[5]

3.10.1 Calculating the Power Spectrum of a Signal

To calculate the power spectrum $P(\omega)$, we need to evaluate the integrals $a_\omega(f)$ and $b_\omega(f)$ for $\omega = 0, 1, 2, \ldots$. Often, the properties of definite integrals of even and odd functions over symmetric intervals can shorten the computations significantly.

3.10.2 The Project: Produce a Spectrum Plot of the Square Wave

We will use numerical integration to compute the spectrum of the square wave

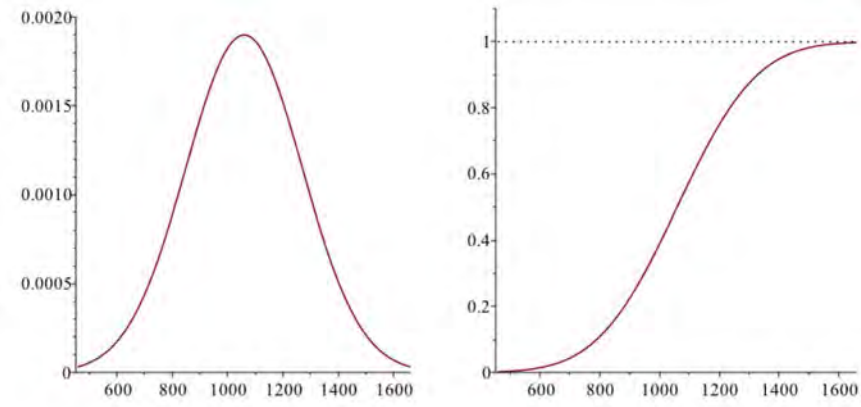
$$SW(t) = \frac{|\sin(3t)|}{\sin(3t)}.$$

1. Choose a method from: *midpoint, trapezoid, Simpson's rule*.
2. Use your method to numerically integrate $P(\omega)$ for $\omega = 1..100$.
3. Use Maple to create a list *Spectrum* of lines $[[k, 0], [k, P(k)]]$ for $k = 1..100$.
4. Graph the list with $plot\big[\{Spectrum\}, thickness = 3\big]$.
5. Change the numerical integration method to either *Gauss quadrature* or *Gauss-Kronrod quadrature*.
6. Recompute the power spectrum and its graph. Do the spectrum values change? Is the procedure faster or slower?

3.10.3 Extra for Experts

1. Show that $a(k) = 0$ for $k \geq 1$. (HINT: *Integrals of even & odd functions.)*
2. Then conclude that $P(k) = |b(k)|$ for $k \geq 1$.
3. Demonstrate that $b(k) = 0$ unless $k = 3 \pmod 6$.

[5] Jean-Baptiste-Joseph Fourier, *Théorie Analytique de la Chaleur*, Firmin Didot, 1822, pg. i.



(a) PDF: *probability density function* (b) CDF: *cumulative distribution function*

Fig. 3.5 The normal distribution for 2019 SAT scores

4. Determine the condition that makes $b(k) = 0$ for $\mathrm{SW}_n(t) = |\sin(nt)|/\sin(nt)$ given some $n \in \mathbb{N}$.

3.11 One Function for All, The Normal Distribution

The *normal distribution* is a surprisingly useful model for a wide range of situations. SAT scores generally fall into a pattern matching a normal curve. The mean and standard deviation for total SAT scores for 2019 are $\mu = 1059$ (from Evidence-Based Reading 531 + Math 528) and $\sigma = 210$, respectively.[6] Define the SAT's *normal probability density function* (also called the *Gaussian distribution*) as

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}.$$

Then the SAT's *normal cumulative distribution function* is

$$N(x) = \int_{-\infty}^{x} p(t)dt.$$

The number of students scoring S is then approximately

$$N(s) = \int_{-\infty}^{s} p(t)\, dt.$$

[6] From the 2019 *SAT Suite of Assessments Annual Report*, The College Board.

3.11.1 The Project

1. Estimate the derivative of $N(x)$ for student scores that are one standard deviation above the mean. Interpret what this values indicates about scores.
2. The minimum score for students scoring in the top 10% is found by solving $0.10 = 1 - N(x)$. Use a root finding method to find x.
3. a. Use a quadrature rule to evaluate $N(1190)$—Appalachian State University's mean score for entering first-year students in 2019. How will you handle the lower limit of infinity? (Hint: $N(400.0) = 0.00085 \approx 0$).
   b. Which percentage group do these students fall into? The top 25%? 30%?
4. Compute $N(x)$ for $x \in [1100, 1500]$ by using the Taylor polynomial approximating $p(x)$ of order

   a. $N = 4$,
   b. $N = 6$,
   c. $N = 10$.

   Describe how well the Taylor polynomials fit using graphs.
5. Estimate the error associated in computing N with each of the Taylor polynomials calculated above.
6. Redo the computations using the most current data and a school near you. Check for data at:

   - SAT Suite of Assessments
   - SAT Suite Data and Reports Archive.

3.12 Monte Carlo Integration

The Mean Value Theorem for Integrals states that, for a continuous function f, there is a point $c \in [a, b]$ such that

$$\int_a^b f(x)\,dx = (b-a) \cdot f(c).$$

This theorem in partnership with the definition of $\overline{f}$, the average value of a function over an interval $[a, b]$,

$$\overline{f} = \frac{1}{b-a}\int_a^b f(x)\,dx,$$

leads to probabilistic methods of integration.

3.12.1 Algebraic: The Definite Integral From an Average

A simple Monte Carlo integration method for computing $\int_a^b f(x)\,dx$ combines the Mean Value Theorem for Integrals with the definition of the average value of a function by taking the average of function values for a large number of points chosen randomly from interval $[a, b]$. The integral's value will be approximated by n random points as

$$\int_a^b f(x)\,dx \approx \frac{b-a}{n}\sum_{i=1}^{n} f(x_i).$$

As n gets large, the error of approximation is $O(1/\sqrt{n})$. This error estimate roughly says that multiplying the number of points by 4 decreases the error by (at least) half.

Let's run through a sample computation with Maple.

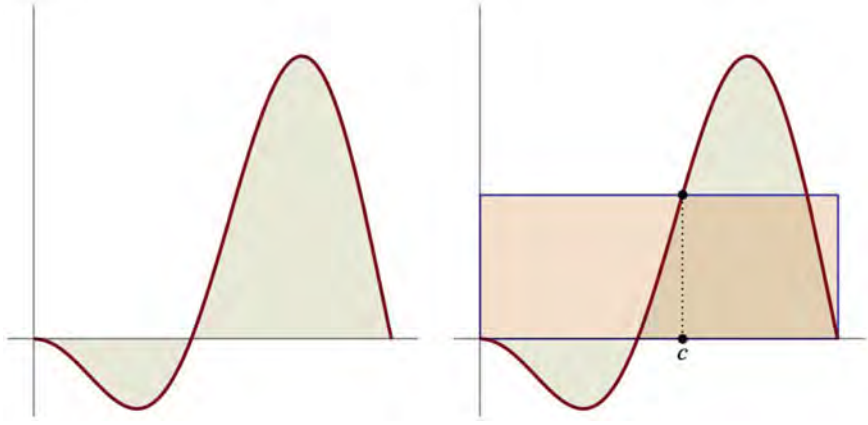
Monte Carlo Integration by Averages

Set $f(x) = x^2/2 - 4x \cdot \sin(2x)$ with $x \in [0, 3.358]$. Estimate

$$\int_0^{3.358} f(x)\,dx.$$

```
> f := x -> 1/2 · x^2 - 4x · sin(2x) :
  a, b := 0, 3.358 :
```

First, we'll seed the random number generator so as to be able to reproduce the sequence



(a) The Area Defined by the Curve (b) Area Generated by One Random Point

Fig. 3.6 Algebraic monte carlo integral approximation

of random values.

```
> randomize(8282623050) :    # Replace with your phone number
```

Generating the random x_is is easy.

```
> N := 10 :
  X := sort([seq(rand(a..b)(), i = 1..N)]) :
```

A little Maple cleverness makes finding the average value extremely easy. A function name followed by a tilde (~) applies the function to all the elements of a list.

```
> (b - a)/N · add(f~(X));
```

$$12.30289947$$

Redoing the computation with $N = 1000$ gives the area estimate 11.94568495. That's not bad—the actual area is 11.98817501.

3.12.2 Geometric: The Definite Integral From Areas

A second simple Monte Carlo integration method for computing $\int_a^b f(x)\,dx$ is based on the interpretation of the integral as the area under curve, and is sometimes called "integration by throwing darts." This method asks whether a random point in a rectangle enclosing the area the function defines is inside the region defined by the function or not. Then the ratio of the "hits," the points inside, to the area of the rectangle gives the estimate of the integral.

In Fig. 3.7, the open gold points are outside the "area under the curve," while the red solid points are inside. Let n be the number of points inside, N be the total number of points generated and the rectangle be $[a, b] \times [c, d]$. Then the ratio of the points inside to the total number of points gives the estimate as

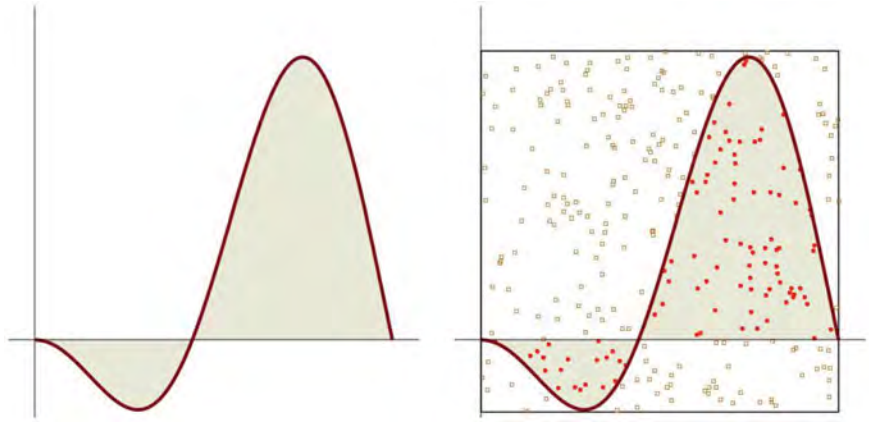
$$\int_a^b f(x)\,dx \approx \frac{n}{N} \cdot (b - a) \cdot (d - c).$$

As n gets large, the error of the geometric method of approximation is also $O(1/\sqrt{n})$.

Let's run through the same example computing again with Maple.

Monte Carlo Integration by Geometry

By looking at a graph, we choose $[a, b] \times [c, d] = [0, 3.358] \times [-3.25, 13]$. (Any rectangle containing the area under the curve is fine—choose convenient values, but the tighter the fit, the better the estimate.) Begin with $N = 300$ (the number of points in Fig. 3.7).



(a) The Area Defined by the Curve (b) Random Points Marking the Area

Fig. 3.7 Geometric monte carlo integral approximation

Set $f(x) = x^2/2 - 4x \cdot \sin(2x)$ with $(x, y) \in [0, 3.358] \times [-3.25, 13]$. Estimate

$$\int_0^{3.358} f(x)\,dx.$$

```
> f := x → 1/2 · x^2 − 4x · sin(2x) :
  a, b := 0, 3.358 :
  c, d := −3.25, 13 :
```

As before, seed the random number generator so as to be able to reproduce the sequence of random values.

```
> randomize(8282623050) :    # Replace with your phone number
```

Generating the set of random (x_i, y_i) points is easy. (Sorting the list isn't really necessary.)

```
> N := 300 :
  Pts := sort([seq([rand(a..b)(), rand(c..d)()], i = 1..N)]) :
```

We need a true/false function that determines if the point is in the area defined by the curve; i.e., is the point between the curve and the x-axis. The test for a point (x, y) is:

$$\text{Does } \underbrace{0 \le y \le f(x)}_{\text{area for positive } f} \text{ or } \underbrace{f(x) \le y \le 0}_{\text{area for negative } f}?$$

In Maple, with $z = [x, y]$,

$$> \mathit{UnderCurve} := z \to (0 \le z_2 \le f(z_1) \textbf{ or } (f(z_1) \le z_2 \le 0)$$

We'll *select* the points that are in the area, then count them with *nops*, (number of operands). Our formula gives

$$> \mathit{XUC} := \mathit{select}(\mathit{UnderCurve}, \mathit{Pts}) :$$
$$\mathit{Estimate} := \frac{\mathit{nops}(\mathit{XUC})}{N} \cdot (b-a) \cdot (d-c);$$
$$\mathit{Estimate} := 17.64349166$$

Try it with 100,000 *points!*

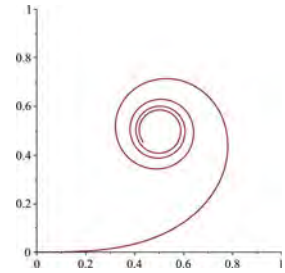
Estimating the error with Monte Carlo integration methods is very subtle. (See, e.g., [4].) As a course approximation, we can use $\varepsilon_N = \mathcal{O}\left(1/\sqrt{N}\right)$.

3.12.3 The Project

Integrate the given functions by both Monte Carlo methods. In each case, compare the answers with those given by Simpson's rule.

1. $f(x) = \sqrt{4 - x^2}$ for $x \in [0, 2]$.
2. $g(x) = e^{-x^2} = \exp(-x^2)$ for $x \in [-1, 1]$.
3. $h(x) = 1/x^2$ for $x \in (0, 1)$.

Fig. 3.8 Transition spiral $\mathcal{T}$



3.13 Computing a Transition Curve

3.13.1 Transition Curves

A *transition curve* is a section of highway or railroad track used to go from a straight section into a curve. A transition curve is designed to prevent sudden changes in lateral acceleration, which would occur with a circular segment, by smoothly transitioning into and out of a curve. The basic transition curve parametric function $\mathcal{T}(t)$ uses the Fresnel sine and cosine integrals

$$\mathrm{C}(t) = \int_0^t \cos\left(\tfrac{1}{2}\pi\, t^2\right)\, dt$$

and

$$\mathrm{S}(t) = \int_0^t \sin\left(\tfrac{1}{2}\pi\, t^2\right)\, dt$$

to give $\mathcal{T}(t) = [\mathrm{C}(t), \mathrm{S}(t)]$. These integrals do not have elementary antiderivatives, so must be evaluated numerically.

A segment $t \in [0, t^*]$ from the spiral $\mathcal{T}$ forms a transition curve which will provide a smooth transition without a sudden change in lateral acceleration. Graph C and S separately to see their respective behaviors.

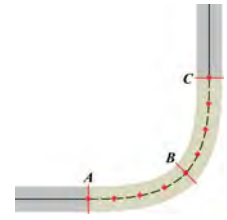
This project will investigate calculating transition curves to lay out a roadbed.

3.13.2 The Project: Computing a Transition Curve for a Roadbed

Two roads are perpendicular and need to be connected by a curve. We'll blend two transition curves to make the road. One curve will start at A and turn towards the north until B. The second transition curve will start at C and turn west meeting the first at B. The two curve's tangents match at B.

We will compute nine waypoints along the centerline using the two transition curves to give data-points to layout the roadbed for surveying and construction. The nine points are spaced evenly along the centerline of the road through the curve. The point A is located at $(0, 0)$. The curve's "minimum radius" (measured at B) is determined by the maximum speed through the curve and is given by the formula

Fig. 3.9 Proposed roadbed



$$R_{min} = \frac{v^2}{127(0.01e_{max} + f_{max})}$$

where the maximum superelevation $e_{max} = 1.0$ and maximum side-friction factor $f_{max} = 1.0$. Our road is in a rural area that has a 100 km/hr speed limit. Point B will be R_{min} from the intersection of perpendiculars at A and C.

Use the information above to determine the coordinates of the nine waypoints on the roadbed.

3.14 Finding the Factorial of $-1/2$

The Gamma function, $\Gamma(x)$, named by Adrien-Marie Legendre, is a very useful extension of the factorial that has no elementary expression for its definition. The function was originally investigated by Daniel Bernoulli in solving the problem of finding a function to interpolate factorials. The expression Bernoulli gave, in a letter to Christian Goldbach in 1729, is

$$\Gamma(x) = \lim_{n\to\infty} \left(n + 1 + \frac{x}{2}\right)^{x-1} \prod_{j=1}^{n} \frac{j+1}{j+x}.$$

A very difficult formula to work with! At Bernoulli's urging Euler sent a letter to Goldbach (about a week after Bernoulli's letter) announcing his solution of the problem with the expression

$$\Gamma(x) = \frac{1 \cdot 2^n}{1+n} \cdot \frac{2^{1-n} \cdot 3^n}{2+n} \cdot \frac{3^{1-n} \cdot 4^n}{3+n} \cdot \frac{4^{1-n} \cdot 5^n}{4+n} \cdots$$

Euler also found several other different formulations of the Gamma function including an integral definition in 1768:

$$\Gamma(x) = \int_0^1 \ln\left(\frac{1}{u}\right)^{x-1} du.$$

3.14.1 The Gamma Function

We will use Adrien-Marie Legendre's formulation of the Γ function [5, pg. 477, 485, 490] resulting from the change of variables $t = \ln(1/u)$,

$$\Gamma(x) = \int_0^\infty t^{x-1} \cdot e^{-t}\, dt.$$

Using *integration by parts* with $u = t^n$ and $v = e^{-t}$ in $\Gamma(n+1)$, it's not hard to see that

$$\Gamma(n+1) = n \cdot \Gamma(n).$$

Combine the recurrence relation above with $\Gamma(1) = 1$ (*Show this!*) to see that

$$\Gamma(n+1) = n!$$

for any integer n. Thus Γ does interpolate the factorial and extend the function to noninteger values except for the negative integers.[7]

3.14.2 The Project

To compute $(-1/2)!$, we'll use $\Gamma(-1/2+1) = \Gamma(1/2)$. Then

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty t^{-1/2} \exp(-t)\, dt.$$

1. Show that $0 \le t^{-1/2} \le 1$ for $t \ge 1$.
2. Conclude that
$$0 < \int_{10}^\infty t^{-1/2} e^{-t}\, dt < \int_{10}^\infty e^{-t}\, dt.$$
3. Analytically evaluate the improper integral $\varepsilon_T = \int_{10}^\infty e^{-t}\, dt$.
4. Approximate $\Gamma(1/2) = (-1/2)!$ by numerically integrating
$$\int_0^{10} t^{-1/2} e^{-t}\, dt$$
and estimating the error. Use
   a. Simpson's rule,
   b. $GK_{7,15}$.
5. Use a calculator to find $(-0.5)!$. How close is the value to your estimates?
6. Square the value you have found. What do you observe?
7. How do $(-0.5)!$ and $(0.5)!$ compare?
8. What is the value of $(-1/2)!$ in Maple?
9. Enter $(-0.25)!$ in Maple and a calculator. What are the results?
10. Graph $\Gamma(x)$ over $[-4.5, 4.5]$. Describe the image.

[7] In fact, Legendre's definition is valid for all complex numbers again except the negative integers.

3.15 Theory: Plugging the Hole in the "Power Rule"

3.15.1 The Integral "Power Rule"

The Power Rule

$$\int x^n \, dx = \frac{x^{n+1}}{n+1} + C \quad \text{for } n \neq -1$$

is the first integration rule learned in calculus classes. This rule, originally called "Cavalieri's quadrature formula" for $n > 0$, was known to Bonaventura Cavalieri 50 years before Newton and Leibniz discovered the Fundamental Theorem of Calculus, the foundational connection between derivatives and integrals. In this project, the Power Rule's missing $(n = -1)$ will be fixed.

3.15.2 The Project

1. Integrate the function $f(x) = x^n$ assuming that $n \neq -1$.
2. Explain why n cannot be -1 in the calculation above.
3. Add the constant $C = -1/(n+1)$ to the integral of the function $f(x) = x^n$ to have

$$\int x^n \, dx = \frac{x^{n+1} - 1}{n+1}.$$

   Define

$$\text{PowerRule}(n) = \frac{x^{n+1} - 1}{n+1}$$

4. Create a table of values for PowerRule(n)

n \ x	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
3.0								
1.0								
0.0								
−0.5								
−0.75								
−0.875								
−0.9375								
−0.96875								

   Maple note: $x = [seq(0.5 \cdot k, k = 1..9)]$ and $n = [seq(-1.0 + 2^{-k}, k = -2..7)]$.
5. Show that both the numerator and denominator of PowerRule(n) go to 0 as n goes to -1.

6. We have a candidate for l'Hôpital's Rule with the variable n. Use l'Hôpital's Rule (differentiating with respect to n) to find the limit

$$\lim_{n=-1} \text{PowerRule}(n).$$

Explain your answer.

References

1. Bauldry, W.C., Bossé, M.J.: A three point construction for the roots of a quartic polynomial
2. Sinnott, R.W.: Virtues of the Haversine. Sky and Telescope **68**(2), 158 (1984)
3. Fornberg, B.: Generation of finite difference formulas on arbitrarily spaced grids. Mathematics of Computation **51**(184), 699–706 (1988)
4. Lazopoulos, A.: Error estimates in monte carlo and quasi-monte carlo integration. Acta Physica Polonica B **35** (2004)
5. de France (Paris), I.: Mémoires de la Classe des Sciences Mathématiques et Physiques de l'Institut de France. Didot (1818)

Further Readings

1. Gauss, C.F.: Methodus nova integralium valores per approximationem inveniendi. apvd Henricvm Dieterich (1815)
2. Kronrod, A.S.: Integration with control of accuracy. SPhD 9, 17 (1964)
3. Maclaurin, C.: A Treatise of Fluxions, vol. 1. Ruddimans (1742)
4. Nahin, P.J.: Inside Interesting Integrals. Springer International Publishing (2020). https://doi.org/10.1007/978-3-030-43788-6. https://doi.org/10.1007%2F978-3-030-43788-6
5. Olde Daalhuis, A.B., et al: NIST Digital Library of Mathematical Functions (2016). https://dlmf.nist.gov
6. Olver, F.W.J., Lozier, D.W., Boisvert, R.F., Clark, C.W.: NIST Handbook of Mathematical Functions (Hardback and CD-ROM). Cambridge University Press (2010)
7. Simpson, T.: A New Treatise of Fluxions: Wherein the Direct and Inverse Method are Demonstrated After a New, Clear, and Concise Manner, with Their Application to Physics and Astronomy: Also the Doctrine of Infinite Series and Reverting Series Universally, are Amply Explained, Fluxionary and Exponential Equations Solved: Together with a Variety of New and Curious Problems. Tho. Gardner (1737)
8. Taylor, B.: Methodus Incrementorum Directa & Inversa. Inny (1717)
9. Villarino, M.B.: Gauss on Gaussian quadrature. The American Mathematical Monthly 127(2), 125–138 (2020). https://doi.org/10.1080/00029890.2020.1680201
10. Zwillinger, D.: The Handbook of Integration. CRC Press (1992)

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

W. C. Bauldry, *Computational Calculus*, Synthesis Lectures on Mathematics & Statistics,
https://doi.org/10.1007/978-3-031-29658-1

Index

A
Asymptotic notation, 6

C
Computer algebra system
  list of, ix

D
Derivative approximation formula
  backward difference, 14
  centered difference, 17
  forward difference, 11
  second derivative
    backward difference, 23
    centered difference, 23
    forward difference, 22
  table of centered difference formulas
    coefficients, 28

E
Error
  round-off error, 12
  truncation error, 12

L
Landau or Big $\mathcal{O}$ notation, 6

M
Maclaurin polynomial, 2
Maple command
  ApproximateInt, 50
  ApproximateIntTutor, 52
  convert to polynomial, 9
  eval, 25
  evalf of Int, 54
  infolevel, 54
  int, numeric, 54
  mtaylor, 8
  nops, 96
  randomize, 57
  select, 96
  solve, 57
    with assumptions, 57
  student package, 9
    Calculus1 subpackage, 9
  taylor, 8
  TaylorApproximationTutor, 9
  unapply, 9
  with, 9
Mathematician
  Archimedes of Syracuse, 31
  Bachmann, Paul, 6
  Bernoulli, Daniel, 98
  Cauchy, Augustin-Louis, 5
  Christoffel, Elwin, 54
  Eudoxus of Cnidas, 31
  Gauss, Johann Carl Friedrich, 54
  Gregory, James, 2
  Hippocrates of Chios, 31
  Jacobi, Carl Gustav Jacob, 59
  Kronrod, Aleksandr S., 60
  L'Hôpital, G.F.A. Marquis de, 6
  Lagrange, Joseph-Louis, 5
  Landau, Edmund, 6
  Legendre, Adrien-Marie, 98

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

W. C. Bauldry, *Computational Calculus*, Synthesis Lectures on Mathematics & Statistics,
https://doi.org/10.1007/978-3-031-29658-1

Lindemann, Carl L. F. von, 31
Maclaurin, Colin, 2
Newton, Isaac, 2
Simpson, Thomas, 47
Stieltjes, Thomas Jan, 61
Taylor, Brook, 2

N
`nDeriv`, calculator function, 25

Q
Quadrature formula
basic methods chart, 51
Gauss-Kronrod $GK_{7,15}$, 61
left endpoint sum, 33
left endpoint sum error, 34
midpoint sum, 39
midpoint sum error, 39
right endpoint sum, 37
right endpoint sum error, 37
Simpson's rule, 48
Simpson's rule error, 49
trapezoid rule, 43
trapezoid rule error, 45

T
Taylor polynomial, 2
examples of, 3
forms of the remainder, 5
Taylor's Theorem, 5
original statement, 28
Transition curve, 97